ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

ADIUVANTIBUS

L. KALMÁR, L. RÉDEI ET K. TANDORI

REDIGIT

B. SZ.-NAGY

TOMUS XXVI

FASC. 1—2

SZEGED, 1965

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

KALMÁR LÁSZLÓ, RÉDEI LÁSZLÓ ÉS TANDORI KÁROLY

KÖZREMŰKÖDÉSÉVEL

SZERKESZTI

SZŐKEFALVI-NAGY BÉLA

26. KÖTET

1—2. FÜZET

SZEGED, 1965. JÚNIUS

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI-INTÉZETE

ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

ADIUVANTIBUS

L. KALMÁR, L. RÉDEI ET K. TANDORI

REDIGIT

B. SZ.-NAGY

TOMUS XXVI

1965

SZEGED, 1965

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

KALMÁR LÁSZLÓ, RÉDEI LÁSZLÓ ÉS TANDORI KÁROLY

KÖZREMŰKÖDÉSÉVEL

SZERKESZTI

SZŐKEFALVI-NAGY BÉLA

26. KÖTET

1965

SZEGED, 1965. DECEMBER

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI-INTÉZETE

INDEX — TARTALOM

TOMUS XXVI — 1965 — 26. KÖTET

Alexits, G., Einige Beiträge zur Approximationstheorie ... 211—224
Alexits, G., und Králik, D., Über die Approximation im starken Sinne ... 93—101
Amemiya, I., and Ando, T., Convergence of random products of contractions in Hilbert space ... 239—244
Ando, T., and Amemiya, I., Convergence of random products of contractions in Hilbert space ... 239—244
Baumslag, G., Kovács, L. G., and Neumann, B. H., On products of normal subgroups 145—148
Brown, A., and Douglas, R. G., On maximum theorems for analytic operator functions 325—328
Brown, A., Halmos, P. R., and Shields, A. L., Cesàro operators ... 125—138
Ceder, J. G., On representing functions by Darboux functions ... 283—288
Douglas, R. G., and Brown, A., On maximum theorems for analytic operator functions 325—328
Fodor, G., and Máté, A., On the structure of set mappings and the existence of free sets 1—8
Foiaș, C., et Sz.-Nagy, B., Sur les contractions de l'espace de Hilbert. X. Contractions similaires à des transformations unitaires ... 79—92
——— Corrections et compléments aux Contractions IX ... 193—196
——— Sur les contractions de l'espace de Hilbert. XI. Transformations unicellulaires 301—345
Freud, G., Approximation of continuous functions on compact metric space by linear methods ... 9—14
Fried, E., Representation of partially ordered groups ... 15—18
Fuchs, L., On partially ordered algebras. II ... 35—42
Гечег, Ф., О композиции автоматов без петель ... 269—272
Гечег, Ф., и Пеак, И., Автоматы с изоморфными полугруппами ... 43—48
Gehér, L., On completely continuous and uniformly bounded operators in l^p spaces .. 31—34
Halmos, P. R., Brown, A., and Shields, A. L., Cesàro operators ... 125—138
Hauschild, K., und Pollák, Gy., Über ein Problem von L. Rédei ... 231—232
Hildebrandt, S., Numerischer Wertebereich und normale Dilatationen ... 187—190
Hosszú, M., Über eine Verallgemeinerung der Distributivitätsgleichung ... 103—106
Jajte, R., General theory of summability. I ... 107—116
Kovács, L. G., and Neumann, B. H., An embedding theorem for some countable groups 139—142
——— On the existence of Baur-soluble groups of arbitrary height ... 143—144
Kovács, L. G., Baumslag, G., and Neumann, B. H., On products of normal subgroups 145—148
Kővary, T., Asymptotic values of entire functions of finite order with density conditions 233—238
Králik, D., und Alexits, G., Über die Approximation im starken Sinne ... 93—102
van Leeuwen, L. C., Über die zulässigen Ideale in Szépschen Ringerweiterungen ... 149—158
Leindler, L., Über Konvergenz- und Summationseigenschaften von Haarschen Reihen 19—30
——— Über verschiedene Konvergenzarten der trigonometrischen Reihen. II (Die Entwicklung der Ableitungen) ... 117—124
——— Über die Unverschärfbarkeit eines Satzes von Orlicz bezüglich vollständige Systeme ... 245—248
Máté, A., and Fodor, G., On the structure of set mappings and the existence of free sets 1—8
Naĭmark, M. A., On unitary group representation in spaces with indefinite metric ... 201—210
Neumann, B. H., and Kovács, L. G., An embedding theorem for some countable groups 139—142
——— On the existence of Baur-soluble groups of arbitrary height ... 143—144
Neumann, B. H., Kovács, L. G., and Baumslag, G., On products of normal subgroups 145—148
Пеак, И., Автоматы и полугруппы. II ... 49—58

Пеак, И., и Гечег, Ф., Автоматы с изоморфными полугруппами .... 43—44
Pollák, Gy., und Hauschild, K., Über ein Problem von L. Rédei .... 231—232
Prohaska, L., Über die Existenz normaler Komplemente zu gewissen Hallgruppen.... 159—162
Rédei, L., Ein Überdeckungssatz für endliche abelsche Gruppen im Zusammenhang mit dem Hauptsatz von Hajós .... 55—62
Rényi, A., On certain representations of real numbers and on sequences of equivalent events .... 63—74
Ruzsa, I., Ein axiomatischer Aufbau eines Systems der deontischen Logik............ 253—268
Sarason, D., On spectral sets having connected complement........................ 289—300
Seibt, H., Zur Quadratbildung von Polynomen.................................. 163—170
Shields, A. L., Brown, A., and Halmos, P. R., Cesàro operators..................... 125—138
Stachó, L., Über ein Problem für Kreisscheibenfamilien.......................... 273—282
Стечкин, С. Б., Неравенства между нормами произвольных производной функции .... 225—230
Sz.-Nagy, B., Positive definite kernels generated by operator-valued analytic functions 191—192
Sz.-Nagy, B., et Foiaş, C., Sur les contractions de l'espace de Hilbert. X. Contractions similaires à des transformations unitaires .... 79—92
——— Corrections et compléments aux Contractions IX .... 193—196
——— Sur les contractions de l'espace de Hilbert. XI. Transformations unicellulaires 301—345
Tandori, K., Über ein Problem von S. B. Stetschkin............................. 75—78
——— Bemerkung zur Konvergenz der Orthogonalreihen.................... 249—252
Weinert, H. J., Zur Theorie der Algebren und monomialen Ringe.................. 171—186

BIBLIOGRAPHIE

Demetrios A. Kappos, Strukturentheorie der Wahrscheinlichkeitsfelder. — G. Scorza Dragoni, Elemente di analisi matematica. Vol. I. Elementi di algebra, Vol. II. La continuità e la differenziabilità, Vol. III. La teoria elementare dell'integrazione. — Á. Császár, Foundations of general topology. — J. Horn und H. Wittich, Gewöhnliche Differentialgleichungen. — M. Barner, Differential- und Integralrechnung, I, Grenzwertbegriff, Differentialgleichung. — E. F. Beckenbach and R. Bellman, Inequalities............................... 197—200
Ewald Burger, Einführung in die Theorie der Spiele. — H. S. Ruse, A. G. Walker and T. J. Villmore, Harmonic spaces. — L. Fuchs, Partially ordered algebraic systems. — L. Fejes Tóth, Reguläre Figuren................. 329—331
Livres reçus par la rédaction 332

KALMÁR LÁSZLÓ

On the structure of set mappings and the existence of free sets

By G. FODOR and A. MÁTÉ in Szeged

To Professor László Kalmár on his 60th birthday

Let E be an infinite set of power m and suppose that to every element x of E there corresponds a subset $S(x)$ of E such that $x \notin S(x)$. Two distinct elements of E, x and y, are called *independent* if $x \notin S(y)$ and $y \notin S(x)$. A subset F of E is called *free* if any two distinct elements of F are independent. *Throughout this paper we assume that m is regular and the power of the set $S(x)$ is $<m$ for every $x \in E$.*

Let us denote by $\mathbf{Pt}(E)$ the set of all subsets of E. Let $\mathbf{S} \subseteqq \mathbf{Pt}(E)$. We say that $\mathbf{S}$ satisfies the ascending chain condition for the ordinal number τ, if there exists no sequence $\{X_\xi\}_{\xi<\tau}$ of type τ of elements of $\mathbf{S}$ such that $X_\vartheta \subset X_\lambda$ holds for every $\vartheta < \lambda < \tau$. Let us denote by $\mathbf{Ac}(E, \tau)$ the set of all sets $\mathbf{S} \subseteqq \mathbf{Pt}(E)$ which satisfy the ascending chain condition for τ. By $S|H$ we denote the set $\{S(x) \cap H : x \in E\}$, where $H \subseteqq E$.

We consider the following two conditions for the set mapping $S(x)$:

(A) *For every $X \in \mathfrak{S} = \{S(x)\}_{x \in E}$, the set $\{Y \in \mathfrak{S} : Y \subseteqq X\}$ is well ordered by $\subseteqq$ and the ordinal type $\xi(X)$ of the set $\{Y \in \mathfrak{S} : Y \subset X\}$ is less than $\omega(m)$. The least upper bound of the ordinal numbers $\xi(X)$, where $X \in \mathfrak{S}$, is $\omega(m)$. The set of the elements $X \in \mathfrak{S}$ for which $\xi(X) = \xi$ has power less than m, and every subset of $\mathfrak{S}$ which is well ordered by $\subseteqq$ has power $<m$.*

(B) *For every decomposition $E = \bigcup_{\eta<\tau} E_\eta$ of E, where $\tau < \omega(m)$, $\overline{\overline{E_\eta}} = m$, and the sets E_η are mutually disjoint, there exists an ordinal number $\xi < \tau$ such that $S|E_\xi \in \mathbf{Ac}(E_\xi, \omega(m))$.*

In this paper we deal with the following question:

Whether or not the condition (A) (or (B)) implies the existence of free subsets of certain cardinalities of E?

In section I we shall prove that the condition (A) does not imply the existence of a free set of power m. In section II we shall prove that the condition (B) implies the existence of a free set of power $\aleph_0$.

We shall use the following notations: For any set $\mathbf{P}$ of sets let

$$\langle \mathbf{P} \rangle = \bigcup_{X \in \mathbf{P}} X.$$

For any cardinal number p we denote by $\omega(p)$ the initial number of p and by p^+ the cardinal number following p immediately. The symbols $\overline{\overline{S}}$ and $\bar{\gamma}$ denote the cardinal numbers of the set S and of the ordinal number γ, respectively. For any element x of E let $S^{-1}(x)=\{y\in E: x\in S(y)\}$ and for any $X\subseteqq E$ let $S(X)=\bigcup_{x\in X} S(x)$ and $S^{-1}(X)=\bigcup_{x\in X} S^{-1}(x)$.

I.

First we consider the following conditions:

(a) *There is a cardinal number $n<m$ such that, for every $x\in E$, $\overline{\overline{S(x)}}<n$.*

(b) *For every $x\in E$, $\overline{\overline{S(x)}}<m$.*

If the condition (a) is satisfied, then there exists a free subset of E with the power m. (See [1].)

On the other hand, it is easy to see that (b) does not imply the existence of an independent pair.

Example: $E=\{\xi:\xi<\omega(m)\}$ and $S(\xi)=\{\eta:\eta<\xi\}$ for every $\xi\in E$.

In this case the set $\mathfrak{S}$ of all sets $S(x)$ forms a well ordered set of type $\omega(m)$ with respect to inclusion $\subseteqq$.

In the following we assume that the condition (b) *holds for the sets $S(x)$, where $x\in E$.*

Now, we consider the following condition for the system $\mathfrak{S}=\{S(x)\}_{x\in E}$:

(c) *For every $X\in\mathfrak{S}$, the set $\{Y\in\mathfrak{S}: Y\subseteqq X\}$ is well ordered by $\subseteqq$ and the ordinal type $\xi(X)$ of the set $\{Y\in\mathfrak{S}: Y\subset X\}$ is less than $\omega(m)$.*

The ordinal number $\xi(X)$ is called the order of the element X. The order of the system $\mathfrak{S}$ is the least upper bound of the orders of its elements.

It is easy to prove that there exists a system $\mathfrak{S}=\{S(x)\}_{x\in E}$ of order 1 such that there exists no free subset of power greater than 2 of E. Let E_1 and E_2 two disjoint subsets of power m of E such that $E=E_1\cup E_2$. Let $\{x_\xi^1\}_{\xi<\omega(m)}$ and $\{x_\xi^2\}_{\xi<\omega(m)}$ be well orderings of E_1 and E_2 respectively. If $x=x_\gamma^1\in E_1$, then let

$$S(x)=\{x_\gamma^2\}\cup\{x_\xi^1\}_{\xi<\gamma}$$

and if $x=x_\gamma^2\in E_2$, then let similarly

$$S(x)=\{x_\gamma^1\}\cup\{x_\xi^2\}_{\xi<\gamma}.$$

It is easy to see that the system $\mathfrak{S}$ of the sets $S(x)$ has the order 1 and there exists no free subset of power greater than 2 of E.

Therefore in the sequel we assume the following condition:

(A) *The system $\mathfrak{S}=\{S(x)\}_{x\in E}$ satisfies the conditions* (b) *and* (c), *the order of $\mathfrak{S}$ is $\omega(m)$, the set of the elements of order ξ of $\mathfrak{S}$ has power less than m, and every subset of $\mathfrak{S}$ which is well ordered by $\subseteqq$ has power $<m$.*

In this section we deal with the following question:

Whether or not the condition (A) *implies the existence of a free subset of power m of E?*

In discussing of this question we shall need another formulation for the properties of the system $\mathfrak{S}$.

Definition 1. 1. The ordered pair $(R,\leqq)$ is said to be a ramification system if R is a set and $\leqq$ is a partial ordering relation defined on the set R satisfying the condition: for every $x\in R$ the set $\{y\in R: y\leqq x\}$ is well ordered by $\leqq$.

Definition 1. 2. The order $\xi(x)$ of an element x of a ramification system $(R, \leqq)$ is the ordinal type of the well ordered set $\{y \in R : y < x\}$.

Definition 1. 3. The order of the system $(R, \leqq)$ is the least upper bound of the orders of its elements.

Definition 1. 4. By a (p, β) ramification system we mean a system of order β, where for each $\xi < \beta$ the set R_ξ of the elements of order ξ has power less than p.

Definition 1. 5. We say that the property $Q(p, \beta)$ holds, if there is a (p, β) ramification system $(R, \leqq)$ of order β such that every subset of R which is well ordered by $\leqq$ has power $<\overline{\beta}$.

It is known that the property $Q(p, \omega(p))$ holds (see [2]), if

(i) $p = n^+$, where n is strongly inaccessible,

(ii) $p = n^+$, where n is regular (assuming the generalized continuum hypothesis),

(iii) p is singular.

We prove now the following

Theorem 1. *The condition* (A) *does not imply the existence of a free set of power* m.

Proof. We shall define a system $\mathfrak{S} = \{S(x)\}_{x \in E}$ which satisfies the condition (A) and for which Theorem 1 is valid, assuming that $Q(m, \omega(m))$ holds (otherwise (A) cannot be satisfied). In this case there exists a ramification system $(R, \leqq)$ satisfying the following conditions:

(1) $$R = \bigcup_{\xi < \omega(m)} R_\xi, \quad R_{\xi_1} \cap R_{\xi_2} = 0 \quad \text{for every} \quad \xi_1 < \xi_2 < \omega(m),$$

where R_ξ is the non-empty set of the elements of order ξ of R,

(2) $$0 < \overline{\overline{R}}_\xi < m$$

for every $\xi < \omega(m)$,

(3) $$\overline{\overline{R}} = m$$

(this follows from (1) and (2)),

(4) $$\overline{\overline{R'}} < m$$

for every subset R' of R which is well ordered by $\leqq$.

Consider now an arbitrary element r of R. We define the set $S(r)$ as follows. There is an ordinal number $\xi < \omega(m)$ such that $r \in R_\xi$. Let

$$S(r) = \bigcup_{\alpha \leqq \xi} R_\alpha - W(r)$$

where $W(r) = \{y \in R : y \leqq r\}$. It is obvious that for every $r \in R$, the power of the set $S(r)$ is $< m$. We prove now

(j) for the system $\mathfrak{S} = \{S(r)\}_{r \in R}$ the condition (A) holds,

(jj) there exists no free set of power m.

Since for $r, s \in R$ the relation $r < s$ holds if and only if the relation $S(r) \subset S(s)$ holds, (j) follows from (1)—(4).

Since r and s are independent if and only if $r < s$ or $s < r$, any free set is ordered, consequently it is well ordered (see 1. 1) and so because of (4) it has power $< m$. Put $E = R$. The theorem is proved.

II.

We assume in the sequel that m is regular and the condition $\overline{\overline{S(x)}} < m$ holds for $x \in E$. In this section we prove the following

Theorem 2. *The condition* (B) *implies the existence of a free set of power $\aleph_0$.*

We need, for the proof of Theorem 2, some lemmas and definitions.

Lemma 2. 1. *Let F be a set of power m, where m is regular and $m \geqq \aleph_0$. Further let $\mathbf{S} \subseteqq \mathbf{Pt}(F)$ such that*

(1) $\mathbf{S} \in \mathbf{Ac}(F, \omega(m))$,

(2) $\overline{\overline{X}} < m$ *for every* $X \in \mathbf{S}$.

Then there exists a subset K of power $< m$ of F such that

(3) $K \nsubseteqq X$ *for every* $X \in \mathbf{S}$.

Proof. Consider the partial ordering of $\mathbf{S}$ with respect to the relation of inclusion. By a theorem of HAUSDORFF [3] there is a maximal ordered subset $\mathbf{P}$ of $\mathbf{S}$. By another theorem of HAUSDORFF [3] $\mathbf{P}$ has a well ordered subset $\mathbf{Q}$ which is confinal to $\mathbf{P}$. It is obvious that $\langle \mathbf{P} \rangle = \langle \mathbf{Q} \rangle$. It follows from (1) that $\overline{\overline{\mathbf{Q}}} < m$. Since m is regular and $\mathbf{Q} \subseteqq \mathbf{S}$, we obtain from (2) that $\overline{\overline{\langle \mathbf{Q} \rangle}} < m$. As $\langle \mathbf{P} \rangle = \langle \mathbf{Q} \rangle$, we have that $\overline{\overline{\langle \mathbf{P} \rangle}} < m$.

Since $\overline{\overline{F}} = m$, $F - \langle \mathbf{P} \rangle \neq 0$. Let x be an arbitrary element of $F - \langle \mathbf{P} \rangle$. Clearly the power of $\langle \mathbf{P} \rangle \cup \{x\}$ is $< m$. By the maximality of $\mathbf{P}$ there exists no set X in $\mathbf{S}$ for which $\langle \mathbf{P} \rangle \cup \{x\} \subseteqq X$ holds. Put $K = \langle \mathbf{P} \rangle \cup \{x\}$. The lemma is proved.

Lemma 2. 2. *The condition* (B) *with $\tau = 1$ implies the existence of a subset K of power $< m$ of E for which $\bigcap_{x \in K} S^{-1}(x) = 0$ holds.*

Proof. It follows from (B) with $\tau = 1$ that $S|E$ satisfies the conditions of Lemma 2. 1 with $E = F$ and $S|E = \mathbf{S}$. Consequently there is a subset K of power $< m$ of E such that $K \nsubseteqq S(x)$ for every $x \in E$, i. e. $\bigcap_{x \in K} S^{-1}(x) = 0$.

Definition 2. 3. Let $\mathfrak{T} = (E, \leqq)$ be an arbitrary partial ordering of the set E. If $x \in E$, then we denote by $\mathfrak{T}(x)$ the set of the minimal elements of the set $\{y \in E: x < y\}$; moreover let $\mathfrak{T}^{(0)}(x) = x$ and $\mathfrak{T}^{(k)}(x) = \mathfrak{T}(\mathfrak{T}^{(k-1)}(x))$ for the natural number k. If $X \subseteqq E$ then we put $\mathfrak{T}(X) = \bigcup_{x \in X} \mathfrak{T}(x)$, $\mathfrak{T}^{(0)}(X) = X$, and $\mathfrak{T}^{(k)}(X) = \mathfrak{T}(\mathfrak{T}^{(k-1)}(X))$.

Definition 2. 4. Let K be the set defined in Lemma 2. 2. The partial ordering $\mathscr{F} = (E, \leqq)$ is said to be free if

(1) $\mathscr{F} = (E, \leqq)$ is a ramification system (defined in 1. 1),

(2) each element of K is minimal in $\mathscr{F}$ and if the element $x \in E - K$ is minimal in $\mathscr{F}$, then it is maximal too.

(3) if $x<y$ in $\mathscr{F}$, then x and y are independent.

Definition 2. 5. The partial ordering $\mathscr{R}=(E,\leqq)$ is said to be regular if
(1) $\mathscr{R}$ is free,
(2) for each $x\in E$, $\overline{\overline{\mathscr{R}(x)}}<m$ (see Definition 2. 3),
(3) for each $x\in E$ for which $\mathscr{R}(x)\neq 0$, $\bigcap_{y\in\mathscr{R}(x)} S^{-1}(y)=0$.

Definition 2. 6. If the partial ordering $\mathscr{R}=(E,\leqq)$ is regular, then the set $X\subseteqq E$ is said to be a path if
(1) X is well ordered by $\leqq$,
(2) if $x\in X$ and $y<x$, then $y\in X$,
(3) $X\cap K\neq 0$ (for K, see Lemma 2. 2 and Definition 2. 4).
Let us denote by $\sigma(\mathscr{R})$ the set of the paths in $\mathscr{R}$, which are maximal with respect to the relation of inclusion $\subseteqq$. Further, for every natural number k, let us denote by $\sigma(\mathscr{R},k)$ the set of the paths of power k in $\mathscr{R}$.

Lemma 2. 7. *Let $\mathscr{R}$ be a regular partial ordering having only paths of finite length. Then*

$$(1)\quad \overline{\overline{\sigma(\mathscr{R})}}<m \quad \text{and} \quad (2)\quad \bigcap_{X\in\sigma(\mathscr{R})} S^{-1}(X)=0.$$

Proof. First we prove the statement (1). It is easy to see by 2. 5/(2), that for every positive integer k

$$\overline{\overline{\sigma(\mathscr{R},k)}}=\overline{\overline{\mathscr{R}^{k-1}(K)}}<m.$$

If $m>\aleph_0$, then by the regularity of m we obtain:

$$\overline{\overline{\sigma(\mathscr{R})}}\leqq\overline{\overline{\bigcup_{1\leqq k<\omega}\sigma(\mathscr{R},k)}}<m.$$

If $m=\aleph_0$, then a simple argument of D. König [4] gives the existence of a positive integer t with $\mathscr{R}^t(K)=0$, and so we have

$$\overline{\overline{\sigma(\mathscr{R})}}=\overline{\overline{\bigcup_{1\leqq k<t}\sigma(\mathscr{R},k)}}<\aleph_0$$

which proves the statement (1).

To prove the statement (2) let y be an arbitrary element of E. It is enough to show that there exists a path $X\in\sigma(\mathscr{R})$ for which $y\notin S^{-1}(X)$. For this purpose let x_0 be an element of K with $y\notin S^{-1}(x_0)$ (such an x_0 clearly exists — see 2. 2). Now suppose that for the positive integer k the path $X_k=\{x_i\}_{i<k}$ has been already defined such that $y\notin S^{-1}(X_k)$. If X_k is no maximal path, then let $x_k\in\mathscr{R}(x_{k-1})$ for which $y\notin S^{-1}(x_k)$ (such an x_k clearly exists — see 2.5/(3)). Since every path in $\mathscr{R}$ is finite, in a finite number of steps we obtain a maximal path $X=X_t=\{x_i\}_{i<t}$ with the positive integer t, such that $y\notin S^{-1}(X)$. The lemma is proved.

Definition 2. 8. Let $\mathfrak{X}$ denote the set of all regular partial orderings $\mathscr{R}$. We define a partial ordering $\mathfrak{M}=(\mathfrak{X},\leqq)$ as follows. Let $\mathscr{R}_1$ and $\mathscr{R}_2$ be two elements of $\mathfrak{X}$, then we put $\mathscr{R}_1\leqq\mathscr{R}_2$ if
(1) the relation $x\leqq y$ in $\mathscr{R}_1$ implies the relation $x\leqq y$ in $\mathscr{R}_2$ and
(2) $\mathscr{R}_1(x)\neq 0$ implies $\mathscr{R}_1(x)=\mathscr{R}_2(x)$.

Lemma 2.9. *Let $\mathfrak{X}'$ be an ordered subset of $\mathfrak{X}$. Then*

(1) $\mathscr{R}' = \bigcup_{\mathscr{R}\in\mathfrak{X}'} \mathscr{R}\in\mathfrak{X}$ *and* (2) $\mathscr{R}\in\mathfrak{X}'$ *implies* $\mathscr{R}\leqq\mathscr{R}'$.

We omit the proof.

It follows from 2.9 with the aid of the Kuratowski—Zorn lemma the following

Lemma 2.10. *The set $\mathfrak{X}$ has a maximal element in $\mathfrak{M}$.*

Lemma 2.11. *If $\mathscr{R}$ is an element of $\mathfrak{X}$ having only paths of finite length, then $\mathscr{R}$ is not maximal in $\mathfrak{M}$.*

Proof. By 2.7/(1) we have $\overline{\overline{\sigma(\mathscr{R})}}<m$. Let

$$X_0, X_1, \dots, X_\xi, \dots \quad (\xi<\tau)$$

be a well ordering of $\overline{\overline{\sigma(\mathscr{R})}}$ for some $\tau<\omega(m)$. As $\overline{\overline{\sigma(\mathscr{R})}}<m$ and for every $\xi<\tau$ $\overline{\overline{X}}_\xi<\aleph_0$, we have $\overline{\overline{\langle\sigma(\mathscr{R})\rangle}}<m$. Thus $\overline{\overline{S(\langle\sigma(\mathscr{R})\rangle)}}<m$ because m is regular and for every $x\in E$, $\overline{\overline{S(x)}}<m$. Put $G = E-\langle\sigma(\mathscr{R})\rangle-S(\langle\sigma(\mathscr{R})\rangle)$. It follows that

(i) $$\overline{\overline{E-G}}<m.$$

Let $$G_\xi = \{x\in G\colon X_\xi\cup\{x\} \text{ is free}\}.$$

Obviously $G_\xi = G-S^{-1}(X_\xi)$. Thus

$$\bigcup_{\xi<\tau} G_\xi = \bigcup_{\xi<\tau}(G-S^{-1}(X_\xi)) = G-\bigcap_{\xi<\tau} S^{-1}(X_\xi) = G-\bigcap_{X\in\sigma(\mathscr{R})} S^{-1}(X).$$

By means of 2.7/(2) we have:

$$\bigcup_{\xi<\tau} G_\xi = G-\bigcap_{X\in\sigma(\mathscr{R})} S^{-1}(X) = G.$$

For every $\xi<\tau$, let

$$H_\xi = G_\xi-\bigcup_{\alpha<\xi} G_\alpha.$$

It is obvious that

(ii) $$\bigcup_{\xi<\tau} H_\xi = \bigcup_{\xi<\tau} G_\xi = G.$$

Let now $F_\xi=H_\xi$ if $\overline{\overline{H}}_\xi=m$ and $F_\xi=0$ if $\overline{\overline{H}}_\xi<m$. In accordance with (i) and (ii) we obtain that

$$\overline{\overline{E-\bigcup_{\alpha<\tau} F_\alpha}}<m.$$

It follows from the condition (B) that there is an $F_\xi\neq0$ such that

$$S|F_\xi\in\mathbf{Ac}(F_\xi, \omega(m)).$$

Therefore we can easily conclude by Lemma 2.1 that there is a set $L\subseteqq F_\xi$ with $\overline{\overline{L}}<m$ such that $L\nsubseteqq S(x)$ for every $x\in E$, i.e. $\bigcap_{y\in L} S^{-1}(y)=0$. Now we define the partial ordering $\mathscr{R}'$ as follows. Let $x\leqq y$ in $\mathscr{R}'$ if $x\in X_\xi$ and $y\in L$, in the other cases

let $u \leqq v$ in $\mathscr{R}'$ if and only if $u \leqq v$ in $\mathscr{R}$. It is easy to see that $\mathscr{R}'$ is regular and $\mathscr{R} < \mathscr{R}'$ in $\mathfrak{M}$. Consequently $\mathscr{R}$ is not maximal. Thus Lemma 2. 11 is proved.

Lemma 2. 12. *There exists a regular partial ordering which has an infinite path X.*

Proof. This follows trivially from 2. 10 and 2. 11.

It follows from the definition of the regular partial ordering (2. 5) that the path X defined in 2. 12 is an infinite free set. Thus Theorem 2 is proved.

References

[1] A. Hajnal, Proof of a conjecture of S. Ruziewicz, *Fundamenta Math.*, **50** (1961—1962), 123—128.
[2] P. Erdős and A. Tarski, On some problems involving inaccessible cardinals, *Essays on the Foundations of Mathematics* (Jerusalem, 1961), 50—82.
[3] F. Hausdorff, Grundzüge einer Theorie der geordneten Mengen, *Math. Annalen*, **65** (1908), 435—505.
[4] D. König, Über eine Schlußweise aus dem Endlichen ins Unendliche, *Acta Sci. Math.*, **3** (1927), 121—130.

(Received December 18, 1964)

Approximation of continuous functions on compact metric space by linear methods

By GÉZA FREUD in Budapest

Dedicated to professor L. Kalmár in occasion of his 60th birthday

§ 1

We refer to the following theorem, due to J. P. NATANSON [2]:

Let $\{K_n(t)\}$ be a sequence of 2π-periodic L-integrable functions, for which the relations

$$\int_{-\pi}^{+\pi} K_n(t)\,dt = 1, \quad \int_{-\pi}^{\pi} |K_n(t)|\,dt = O(1)$$

and

$$\int_{-\pi}^{+\pi} |tK_n(t)|\,dt = O(\lambda_n) \qquad (\lambda_n \downarrow 0)$$

are satisfied, and with the aid of $\{K_n(t)\}$ define for arbitrary 2π-periodic continuous functions $f(t)$ the sequence of linear transformations

$$A_n(f; x) = \int_{-\pi}^{\pi} f(t+x)\,K_n(t)\,dt.$$

Then

$$A_n(f; t) - f(t) = O(1)\cdot\omega(f; \lambda_n),$$

where

$$\omega(f; \delta) = \max_{\substack{|h| \leqq \delta \\ x \in (-\pi, +\pi)}} |f(x+h) - f(x)|$$

is the continuity modulus of $f(x)$.

The aim of this paper is to extend this theorem to a rather general case. An example, where the generalized theorem is needed, is contained in § 5.

§ 2

Let K be a compact metric space, with the distance function $\varrho(x, y)$ $(x, y \in K)$, let further C_K be the space of real valued continuous functions $f(x)$ over K with the usual norm

$$\|f(x)\| = \operatorname*{Max}_{x\in K} |f(x)|. \tag{1}$$

Let $\alpha(f)$ be a bounded linear functional over C_K,

$$\sup_{\|f\|\leqq 1} |\alpha(f)| = A \tag{2}$$

and

$$\sup |\alpha\{f(x)\varrho(x,\xi)\}| = A_\xi. \tag{3}$$

Theorem 1. *Let $\varphi(\delta)$ be a not decreasing function with $\varphi(0)=0$ and*

$$\varphi(2\delta)\leqq 2\varphi(\delta) \qquad (\delta>0). \tag{4}$$

Then for every $f\in C_K$ the condition

$$|f(x)-f(\xi)| \leqq \varphi\{\varrho(x,\xi)\} \qquad (x\in K,\ \xi \text{ fixed}) \tag{5}$$

implies for every $\sigma>0$

$$|\alpha(f)-f(\xi)\alpha(1)| \leqq (A+3\sigma^{-1}A_\xi)\varphi(\sigma).^{1)} \tag{6}$$

Before proving our theorem, we deduce some of its consequences, the proof itself is postponed to § 3. We call K convex if for every pair $x_1, x_2\in K$ there is at least one $x_{12}\in K$ such that

$$\varrho(x_i, x_{12}) = \tfrac{1}{2}\varrho(x_1,x_2) \qquad (i=1,2). \tag{7}$$

For convex K the modulus of continuity

$$\omega(f;\delta) = \operatorname*{Max}_{\substack{x_1,x_2\in K\\ \varrho(x_1,x_2)\leqq\delta}} |f(x_1)-f(x_2)|$$

of an arbitrary function $f(x)$ satisfies

$$\omega(f;2\delta)\leqq 2\omega(f;\delta).$$

Let[2])

$$A^{(1)} = \sup_{\xi\in K} A_\xi.$$

Putting $\varphi=\omega$ we obtain from (6)

$$|\alpha(f)-f(\xi)\alpha(1)| \leqq (A+3\sigma^{-1}A^{(1)})\omega(f;\sigma) \tag{6a}$$

for every $\sigma>0$.

---

[1]) We use the notation $\alpha(1)=\alpha(f_0)$, $f_0\equiv 1$.

[2]) As a consequence of the compactness of K

$$\sup_{x_1,x_2\in K} \varrho(x_1,x_2) = R<\infty$$

so that $A_\xi\leqq RA$. From this we conclude $A^{(1)}<\infty$.

§ 3

Let K be convex and let A be a bounded linear transformation of C_K into itself, transforming $f \in C_K$ into $A(f; x) \in C_K$ $(x \in K)$, with the norm

$$\sup_{\|f\| \leqq 1} \|A(f)\| = \|A\|.$$

For each fixed $\xi \in K$ we consider the linear transformation

$$A_\xi(f) = A\{\varrho(x, \xi) f(x)\}$$

and set[3])

$$A^{(1)} = \sup_{\xi \in K} \|A_\xi\|.$$

Putting $\alpha(f) = A(f; \xi)$, $A \leqq \|A\|$, $A^{(1)} \leqq A^{(1)}$ in (6a), we obtain

(6b) $$|A(f; \xi) - A(1; \xi) f(\xi)| \leqq (\|A\| + 3A^{(1)}\sigma^{-1})\omega(f; \sigma)$$

for every $\sigma > 0$ and $\xi \in K$.

Now let us consider a sequence $\{A_n\}$ of bounded linear transformations over C_K, such that

(8) $$\|A_n\| = O(1) \quad \text{and} \quad A_n^{(1)} = O(\lambda_n),$$

where $\lambda_n \downarrow 0$.

Substituting $A = A_n$, $\sigma = \lambda_n$ in (6b), we obtain[4])

$$A_n(f; \xi) - f(\xi) A_n(1; \xi) = O\{\omega(f; \lambda_n)\}$$

and the constant in the O-estimate does not depend on the choice of $f \in C_K$ and $\xi \in K$.

This gives the announced generalization of NATANSON's theorem:

Theorem 2. *Let K be convex, and let the sequence $\{A_n\}$ of linear transformations over C_K satisfy* (8). *Then*

(9) $$|A_n(f; \xi) - f(\xi) A_n(1; \xi)| \leqq K_1 \omega(f; \lambda_n)$$

where K_1 is neither depending on ξ nor on the choice of $f \in C_K$.

§ 4

We turn to the proof of Theorem I.

Lemma 1. *For every $\sigma > 0$ and $\vartheta > 1$ we have*

(10) $$\varphi(\vartheta\sigma) < 2\vartheta\varphi(\sigma).$$

Proof. From (4) it follows by induction

$$\varphi(2^m\delta) \leqq 2^m \varphi(\delta) \qquad (m = 1, 2, \ldots).$$

---

[3]) From the inequality $\|A_\xi\| \leqq R\|A\|$ (see [1])) we conclude $A^{(1)} < \infty$.

[4]) We use the notations $A_n(1; \xi) = A_n(f_0; \xi)$, $f_0 \equiv 1$.

Let now m be the integer for which

$$2^m < \vartheta \leqq 2^{m+1}.$$

Since $\varphi(\delta)$ is monotone, we obtain

$$\varphi(\vartheta\sigma) \leqq \varphi(2^{m+1}\sigma) \leqq 2^{m+1}\varphi(\sigma) < 2\vartheta\varphi(\sigma). \qquad \text{Q. e. d.}$$

Lemma 2. *If, for a fixed $\xi \in K$, (5) and (4) are satisfied, then for arbitrary $\sigma > 0$ there is an $f_1 \in C_K$ and an $f_2 \in C_K$ such that*

$$f(x) = f(\xi) + [f_1(x) + 3\sigma^{-1}\varrho(x, \xi) f_2(x)]\,\varphi(\sigma), \tag{11}$$

where

$$\|f_1\| \leqq 1 \quad \textit{and} \quad \|f_2\| \leqq 1. \tag{12}$$

Proof. We consider the function $F(x) = f(x) - f(\xi)$ on the closed set $\Sigma = \{x : \varrho(x, \xi) \leqq \sigma\}$. According to the theorem of Tietze F can be extended as a continuous function to K, so that

$$\max_{x \in K} |F(x)| = \max_{x \in \Sigma} |F(x)| \leqq \varphi(\sigma).$$

We put $F(x) = f_1(x)\varphi(\sigma)$, $f_1 \in C_K$, $\|f_1\| \leqq 1$, and define $f_2 \in C_K$ by (11). Then $f_2(x) = 0$ for $x \in \Sigma$, and for $x \notin \Sigma$ (i. e. $\varrho(x, \xi) > \sigma$) we have by Lemma 1 with $\vartheta = \sigma^{-1}\varrho(x, \xi)$

$$|f(x) - f(\xi)| \leqq \varphi\{\varrho(x, \xi)\} \leqq 2\sigma^{-1}\varrho(x, \xi)\,\varphi(\sigma).$$

For $x \notin \Sigma$ we have

$$|F(x)| \leqq \varphi(\sigma) \leqq \sigma^{-1}\varrho(x, \xi)\varphi(\sigma),$$

so that (11) gives

$$|f_2(x)| \leqq 1, \quad x \notin \Sigma.$$

For $x \in \Sigma$ we had $f_2(x) = 0$, so that

$$\|f_2\| \leqq 1. \qquad \text{Q. e. d.}$$

Proof of Theorem 1. From the representation (11) we conclude

$$\alpha(f) - f(\xi)\alpha(1) = [\alpha(f_1) + 3\sigma^{-1}\alpha\{\varrho(x, \xi) f_2(x)\}]\varphi(\sigma),$$

hence by (12), (2) and (3) we obtain

$$|\alpha(f_1)| \leqq A$$

and

$$|\alpha\{\varrho(x, \xi) f_2(x)\}| \leqq A_\xi,$$

so that

$$|\alpha(f) - f(\xi)\alpha(1)| \leqq (A + 3\sigma^{-1}A_\xi)\varphi(\sigma). \qquad \text{Q. e. d.}$$

§ 5

We mention a typical application of our theorem. In [1] we applied the following lemma: We consider a matrix of real functions, $\{\varphi_{kn}(x)\}$ $(k=1, 2, \ldots, n;$ $n=1, 2, \ldots)$ defined over a finite interval $[a, b]$, $a<0<b$. For $x \in [a, b]$ let

$$\sum_{k=1}^{n} \varphi_{kn}(x) = 1 + O\left(\frac{1}{n}\right), \tag{13}$$

$$\sum_{k=1}^{n} |x - x_{kn}|\,|\varphi_{kn}(x)| = O\left(\frac{1}{n}\right), \tag{14}$$

and

$$\sum_{k=1}^{n} |\varphi_{kn}(x)| = O(1). \tag{15}$$

Then for every $g \in C[a, b]$

$$g(0) + \sum_{k=1}^{n} \varphi_{kn}(x)[g(x_{kn}) - g(0)] - g(x) = O(1)\,\omega\left(g; \frac{1}{4n}\right).$$

Setting

$$\Lambda_n(f; x) = \frac{\sum_{k=1}^{n} \varphi_{kn}(x) f(x_{kn})}{\sum_{k=1}^{n} \varphi_{kn}(x)},$$

$f(x) = g(x) - g(0)$, and $\lambda_n = \dfrac{1}{4n}$, the conditions of Theorem 2 are satisfied and we obtain as its conclusion

$$\Lambda_n(f; x) - f(x) = O(1)\cdot\omega\left(f; \frac{1}{4n}\right),$$

i. e.

$$g(0) + \sum_{k=1}^{n} \varphi_{kn}(x)[g(x_{kn}) - g(0)] - g(x) =$$

$$= O(1)\,\omega\left(g; \frac{1}{n}\right) + \max|g(x) - g(0)|\,O\left(\frac{1}{n}\right) = O(1)\,\omega\left(g; \frac{1}{4n}\right),$$

so that this lemma appears to be a consequence of Theorem 2, though it would not follow from NATANSON's theorem.

References

[1] G. FREUD, Über ein Jacksonsches Interpolationsverfahren, *Über Approximationstheorie*, ISNM, 5 (Basel—Stuttgart, 1964), 227—232.

[2] И. П. Натансон, О точности представления непрерывных периодических функций сингулярными интегралами, *Доклады Акад. Наук СССР*, **73** (1950), 273—276.

(Received July 16, 1964)

Representation of partially ordered groups

By E. FRIED in Budapest

To Professor L. Kalmár on his 60th birthday

A well-known theorem of CAYLEY states that groups are representable as permutation groups. In his paper [3] Ch. HOLLAND has dealt with a similar problem and proved that each lattice-ordered group is isomorphic to a group of monotone permutations of a suitable fully ordered set. Here we shall consider two questions connected with the above ones. The first one is similar to CAYLEY's theorem and is concerned with partially ordered groups, while the second one is connected with the paper of HOLLAND and characterizes the partially ordered groups that can be regarded as subgroups of the lattice-ordered group of the monotone permutations of some fully ordered set.

The fundamental concepts concerning partially ordered groups can be found in the book of L. FUCHS [2].

The following result is well known.

Lemma. *A group Σ of permutations of a fully ordered set S is partially left-ordered*[1]) *by the rule*[2])

$$\alpha \leqq \beta \quad (\alpha, \beta \text{ in } \Sigma), \quad \textit{if } u^\alpha \leqq u^\beta \textit{ for each } u \in S.$$

Before proving the converse of this lemma we shall consider the connection of S and Σ in more detail.

We define a new partial order $\leqq'$ on the set S in the following way:

Let $a \leqq' b$ (a, b in S) if there exist a u in S and α, β in Σ such that $a = u^\alpha$, $b = u^\beta$ and $\alpha \leqq \beta$. We denote the set S partially ordered under $\leqq'$ by S'. The order of S is obviously an extension of the order of S'. Indeed, $a \leqq' b$ (a, b in S) implies by definition $a \leqq b$.

We verify that the relation $\leqq'$ is actually a partial order. We get reflexivity by the choice $\alpha = \beta = \varepsilon$ (the unity of Σ). Antisimmetry is clear because S is an extension of S'. Finally, we obtain transitivity in the following way: $a \leqq' b$ and $b \leqq' c$ imply the existence of u, v in S and $\alpha \leqq \beta$, $\gamma \leqq \delta$ in Σ with the properties $a = u^\alpha$, $b = u^\beta = v^\gamma$, $c = v^\delta$. From the partial left-order it follows $\beta^{-1}\alpha \leqq \varepsilon \leqq \gamma^{-1}\delta$, that is, $a = u^\alpha = b^{\beta^{-1}\alpha} \leqq' \leqq' b^{\gamma^{-1}\delta} = v^\delta = c$.

---

[1]) A group Σ is partially left-ordered if there is a partial order $\leqq$ in the set Σ with the property: $\alpha \leqq \beta$ implies $\gamma\alpha \leqq \gamma\beta$ for all $\alpha, \beta, \gamma \in \Sigma$. We can define partial right-order similarly. A group is partially ordered if it is both partially left- and right-ordered.

[2]) We denote by u^α the image of the element $u \in S$ under the mapping $\alpha \in \Sigma$.

Thus we have the following

Theorem 1. *S' is a partially ordered set.*

The connection between Σ and S' is shown by

Theorem 2. *Σ is a partially ordered group if and only if its elements are monotone permutations of S'.*

Proof. Let first Σ be a partially ordered group, and let $a \leqq' b$. Then $a = u^\alpha$, $b = u^\beta$ for a suitable $u \in S$ and for some $\alpha \leqq \beta$ in Σ. By two-sided ordering, for every $\sigma \in \Sigma$, $\alpha\sigma \leqq \beta\sigma$ holds, and therefore $a^\sigma = u^{\alpha\sigma} \leqq' u^{\beta\sigma} = b^\sigma$, as asserted.

Next, let the elements of Σ be monotone permutations of S', and let $\alpha \leqq \beta$ and σ in Σ. From $\alpha \leqq \beta$ we obtain for every $u \in S'$ the relation $u^\alpha \leqq' u^\beta$, and hence because of the monotonicity $u^{\alpha\sigma} \leqq' u^{\beta\sigma}$. On using the fact that S is an extension of S', we get $u^{\alpha\sigma} \leqq u^{\beta\sigma}$ for each $u \in S$, that is, $\alpha\sigma \leqq \beta\sigma$.

Clearly, S is the union of the disjoint sets S_ν which are the domains of intransitivity under Σ.

Let Σ_ν be the partially left-ordered group of the permutations induced by the elements of Σ on S_ν. It is easy to see that Σ is a subdirect product of the partially left-ordered groups Σ_ν. (See COHN [1][3].)

Now, let $\bar{\Sigma}$ be a subgroup of Σ. We can define $\overline{S'}$ and $\overline{S_\nu}$ similarly as we did S' and S_ν. Obviously, each S_ν is the union of some $\overline{S_\mu}$, and it is easy to see that the order of S' is an extension of the order of $\overline{S'}$. It can be proved without difficulty that if all positive elements of Σ are in $\bar{\Sigma}$ then $\overline{S'} = S'$.

Theorem 3. *Let $\bar{\Sigma}$ be a normal subgroup of Σ. Then the $\overline{S_\nu}$'s which are subsets of the same S_μ are domains of imprimitivity of Σ.*

Proof. Let $\overline{S_1}$, $\overline{S_2}$ belong to S_1, $a, b \in \overline{S_1}$ and $c \in \overline{S_2}$. Then $b = a^\varphi$ and $c = a^\alpha$, where $\varphi \in \bar{\Sigma}$ and $\alpha \in \Sigma$. Hence $b^\alpha = a^{\varphi\alpha} = a^{\alpha\alpha^{-1}\varphi\alpha} = c^{\alpha^{-1}\varphi\alpha}$. Now, $\alpha^{-1}\varphi\alpha \in \bar{\Sigma}$ because $\bar{\Sigma}$ is a normal subgroup of Σ, thus $b^\alpha \in \overline{S_2}$.

The following result gives information about the representability of partially left-ordered groups.

Theorem 4. *If Σ is a partially left-ordered group then there exists a fully ordered set S such that the partially left-ordered group of the permutations of S contains a subgroup o-isomorphic to Σ.*

Proof. Let Σ be a partially left-ordered group and let Σ_0 denote the underlying partially ordered set (i. e. in Σ_0 we disregard from the group operation). It is known (see SZPILRAJN [4]) that the order of Σ_0 is the intersection of orders of some fully ordered sets Σ_ν. We can suppose that the set of indices ν is ordered. Finally, let S be the union of the disjoint sets Σ'_ν such that Σ'_ν is order-isomorphic to Σ_ν. We denote by u_ν the element of Σ'_ν corresponding to the element u of Σ_0. We define a full order in S by putting $u_\nu \leqq v_\mu$ if either ν precedes μ in the ordering of the indices or if $\nu = \mu$ and $u_\nu \leqq v_\nu$ in Σ'_ν.

Now we define, for each $a \in \Sigma$, a mapping σ_a of S such that $u_\nu^{\sigma_a} = (ua)_\nu$ for each ν. Obviously, σ_a is a permutation of S. The mapping $a \to \sigma_a$ is a one-to-one corres-

[3]) COHN considers only the case of complete direct product.

pondence, for if $\sigma_a=\sigma_b$, then $a_\nu=e_\nu^{\sigma_a}=e_\nu^{\sigma_b}=b_\nu$ for every ν, that is, $a=b$. From $u_\nu^{\sigma_a\sigma_b}=(ua)_\nu^{\sigma_b}=(uab)_\nu=u_\nu^{\sigma_{ab}}$ it follows that the mapping $a\to\sigma_a$ is an isomorphism. Let $a\leqq b$. Then $ua\leqq ub$, and therefore $u_\nu^{\sigma_a}=(ua)_\nu\leqq(ub)_\nu=u_\nu^{\sigma_b}$. Thus $\sigma_a\leqq\sigma_b$, and the mapping $a\to\sigma_a$ is an o-preserving isomorphism. Let finally σ_a be a positive element, that is, $e_\nu\leqq e_\nu^{\sigma_a}=(ea)_\nu=a_\nu$ for all ν. Since Σ_0 is the intersection of the Σ_ν, we get $e\leqq a$, showing that $a\to\sigma_a$ is an o-isomorphism.

Next we shall consider partially ordered groups which are representable by monotone permutations of a fully ordered set. The monotone permutations of a fully ordered set S, as is well known, form a lattice-ordered group under the usual ordering (see: Lemma). HOLLAND proved in [3] that to each lattice-ordered group Σ there exists a fully ordered set S such that the group of all monotone permutations of S, under the usual ordering, has a subgroup o-isomorphic to Σ. We get the obvious

Theorem 5. *A partially ordered group is representable by monotone permutations of a fully ordered set if and only if it is isomorphic to a subgroup of a lattice-ordered group.*

Let the elements of a group Σ operate on a set S. We call an ordering $\leqq$ of S a Σ-ordering if, for each $\alpha\in\Sigma$, $a\leqq b$ implies $a^\alpha\leqq b^\alpha$ $(a, b\in S)$. Let, for example, Σ be a partially ordered group and P a convex subgroup of Σ. The set of the right cosets of P is in the induced partial order ($P\alpha\leqq P\beta$ if there are elements $\gamma\in P\alpha$, $\delta\in P\beta$ such that $\gamma\leqq\delta$) a Σ-ordered set by the definition $(P\alpha)^\sigma=P\alpha\sigma$. If the ordering of the Σ-ordered set of the right cosets of a convex subgroup P can be extended to a full Σ-ordering, we call P an *admissible* subgroup.

Remark. A convex subgroup P of a commutative group Σ is admissible if and only if the induced order of the group Σ/P is extendible to a full order, that is, Σ/P is torsionfree. Let namely $\leqq$ be the induced order in Σ/P and $\leqq'$ a full order of Σ/P which is an extension of $\leqq$. Let $P\alpha\leqq' P\beta$ and τ an element of Σ. Because of $P\varrho P\sigma=P\varrho\sigma$, $P\alpha P\tau\leqq' P\beta P\tau$ and $P\alpha\tau\leqq' P\beta\tau$ are equivalent. The commutativity of Σ completes the proof.

Example. We construct a group containing a subgroup which is not admissible. Let the elements of Σ be the numbers of the form $\pm 2^\alpha$ and the elements of P the numbers of the form 2^α, where α are integers, and the group operation is multiplication of numbers. Let the positive elements be the elements $\geqq 1$; then P is convex and there are two cosets of P, namely P and $-P$. P can not be an admissible subgroup because $P(-1)=-P$ and $-P(-1)=P$.

Theorem 6. *A partially ordered group Σ is o-isomorphic to a subgroup of the lattice-ordered group of all monotone permutations of a fully ordered set S if and only if there exists a set Λ of admissible subgroups P of Σ such that to each non-positive α in Σ there is some P in Λ satisfying $P\alpha<P$ in the induced ordering*[4]).

Proof. Let first Σ be o-isomorphic to a subgroup of the partially ordered group of all monotone permutations of a fully ordered set S. We can assume, on the basis of this o-isomorphism, that the elements of Σ are monotone permutations of S. Let $a\in S$, and define P_a to consist of all $\alpha\in\Sigma$ such that $a^\alpha=a$. Obviously, for each $a\in S$, P_a is a convex subgroup. There will not be any ambiguity in denoting

[4]) The last assumption implies that Σ is directed.

the orders of S, Σ and the set of cosets of P_a by the same sign $\leqq$. Let us fix an a in S. $a^\alpha = a^\beta$ is equivalent to $a^{\alpha\beta^{-1}} = a$, further to $\alpha\beta^{-1} \in P_a$, that is, $a^\alpha = a^\beta$ if and only if α and β are in the same right coset of P_a. Now, we order the cosets of P_a by setting $P_a\alpha \leqq P_a\beta$ if and only if $a^\alpha \leqq a^\beta$. This relation is, obviously, independent of the representation of the cosets. It is also clear that this relation is an order, moreover a full order, because the elements a^α are in the fully ordered set S. Now, $P_a\alpha \leqq P_a\beta$, i. e., $a^\alpha \leqq a^\beta$ implies $a^{\alpha\sigma} \leqq a^{\beta\sigma}$, i. e., $P_a\alpha\sigma \leqq P_a\beta\sigma$. Therefore P_a is admissible.

Finally, let α be a non-positive element in Σ. From the non-positivity we infer the existence of an $a \in S$ such that $a^\alpha < a$, because from $a^\alpha \geqq a$, for each $a \in S$, the positivity of α would follow. From $a^\alpha < a$ it follows $P_a\alpha < P_a$; this means that the subgroup P_a has the property required in theorem 6.

Conversely, let Λ be a set of admissible subgroups P_ν, such that to each non-positive $\alpha \in \Sigma$ there is some P_ν in Λ with the property $P_\nu\alpha < P_\nu$. We can suppose Λ to be fully ordered. Let S_ν denote the set of the right cosets of P_ν; S_ν is a fully Σ-ordered set. We denote the orders in S_ν and in Λ by $\leqq$. Let finally S be the union of S_ν with the following full order: $P_\nu\alpha < P_\mu\beta$ if either $\nu < \mu$ in Λ or if $\nu = \mu$ and $P_\nu\alpha < P_\nu\beta$ in S_ν. We define, for each $\sigma \in \Sigma$, the mapping $P_\nu\alpha \to (P_\nu\alpha)^\sigma = P_\nu\alpha\sigma$ of S into itself. Obviously, these are mappings of S onto S, moreover they are one-to-one. These permutations are monotone. For, let $P_\nu\alpha < P_\mu\beta$. If $\nu < \mu$ in Λ, then $P_\nu\alpha\sigma < P_\mu\beta\sigma$ for each $\sigma \in \Sigma$. If $P_\nu\alpha < P_\nu\beta$ in S_ν, then $P_\nu\alpha\sigma < P_\nu\beta\sigma$, because S_ν is Σ-ordered.

Now, to $\sigma \in \Sigma$ we make correspond the monotone permutation $P_\nu\alpha \to P_\nu\alpha\sigma$ of S. Then Σ will be isomorphic to a subgroup of the monotone permutations of S. We have to show that this isomorphism is an o-isomorphism. It is enough to prove that a monotone permutation $P_\nu\alpha \to P_\nu\alpha\sigma$ is positive if and only if σ is a positive element of Σ. Let σ be positive. Then $\alpha < \alpha\sigma$ for each α in Σ and $P_\nu\alpha < P_\nu\alpha\sigma = (P_\nu\alpha)^\sigma$ for all P_ν in Λ and α in Σ. Now, let σ be non-positive. By hypothesis there exists a subgroup P_ν in Λ such that $P_\nu\sigma < P_\nu$, and so the mapping $P_\nu\alpha \to P_\nu\alpha\sigma$ is again non-positive.

Remark. The intersection of all P_ν in Λ is the unity of Σ. In order to prove this proposition, it is enough to show that for each $\alpha \neq \varepsilon$ in Σ there is a P_ν in Λ such that $\alpha \notin P_\nu$. If α is non-positive then there is a P_ν in Λ such that $P_\nu\alpha < P_\nu$ and so $\alpha \notin P_\nu$. If $\alpha > \varepsilon$, then α^{-1} is non-positive, and if $\alpha^{-1} \notin P_\nu$ then $\alpha \notin P_\nu$ too. If we replace the condition "for each non-positive $\alpha \in \Sigma$ there exists a $P \in \Lambda$ such that $P\alpha < P$ holds" by "the intersection of all $P_\nu \in \Lambda$ is the unity of Σ", we are able to prove only the existence of an o-preserving isomorphism.

References

[1] P. M. Cohn, Groups of order automorphisms of ordered sets, *Mathematika*, **4** (1957), 41–50.
[2] L. Fuchs, *Partially ordered algebraic systems* (Oxford, 1963).
[3] Ch. Holland, The lattice-ordered group of automorphisms of an ordered set, *Michigan Math. J.*, **10** (1963), 399–408.
[4] E. Szpilrajn, Sur l'extension de l'ordre partiel, *Fund. Math.*, **16** (1930), 386–389.

(Received November 6, 1964)

Über Konvergenz- und Summationseigenschaften von Haarschen Reihen

Von LÁSZLÓ LEINDLER in Szeged

Herrn Professor László Kalmár zum 60. Geburtstag gewidmet

Einleitung

Im Falle $\Sigma a_n^2 < \infty$ konvergiert die Haarsche Reihe[1])

$$\sum_{m=0}^{\infty} \sum_{k=1}^{2^m} b_m^{(k)} \chi_m^{(k)}(x) = \sum_{n=0}^{\infty} a_n \chi_n(x) \tag{1}$$

bekanntlich fast überall. Für den Fall $\Sigma a_n^2 = \infty$ gelten dagegen die folgenden Sätze:

Satz I. *Ist* $\sum_{n=0}^{\infty} a_n^2 = \infty$ *und ist die Folge* $\{|a_n|\}$ *monoton nichtwachsend, so divergiert die Haarsche Reihe* (1) *fast überall.*[1a])

Satz II. *Ist* $\sum_{n=0}^{\infty} a_n^2 = \infty$, *so existiert ein System* $\{h_n(x)\}$ *vom Haarschen Typ*[2]), *für welches die Reihe*

$$\sum_{n=0}^{\infty} a_n h_n(x)$$

fast überall divergiert.

Ähnliche Behauptungen gelten für das Rademachersche System; man bemerke aber, daß das Haarsche System „besser" als das Rademachersche System ist, da die Lebesgueschen Funktionen des Haarschen Systems beschränkt sind. Wenn wir die Beschränktheit der Lebesgueschen Funktionen in Betracht nehmen, können wir einen Satz von K. TANDORI [7] auf Grund des Satzes I folgenderweise verallgemeinern:

Satz III. *Ist* $\{\lambda_n\}$ *eine positive, nichtabnehmende Zahlenfolge mit* $\lambda_n = O(\log^2 n)$, *so existiert ein im Intervall* (a, b) *orthonormiertes Funktionensystem* $\{f_n(x)\}$ *mit den*

[1a]) Während der Drucklegung unserer Arbeit erschien ein Aufsatz von P. L. Uljanov (П. Л. Ульянов, О рядах по системе Хаара с монотонными коэффициентами, *Известия АН СССР*, **28** (1964), 925–950) in welchem mehrere Sätze über das Haarsche System bewiesen sind, die unseren Satz I und teils den Satz IV enthalten.

[1]) Siehe [1] S. 48.

[2]) Ein im Intervall (0, 1) orthonormiertes Funktionensystem $\{h_n(x)\}$ ist vom Haarschen Typ, wenn für jedes Indexpaar m, n mit $m \neq n$ und $2^k < m, n \leq 2^{k+1}$ $(k = 0, 1, \ldots)$ gilt: $h_n(x) h_m(x) = 0$ überall in (0, 1).

folgenden Eigenschaften: es gilt in (a, b) $L_n(x) = O(\lambda_n)$[3]) *und für jede positive Zahlenfolge* $\{w_n\}$ *mit* $w_n = o(\lambda_n)$ *gibt es eine Koeffizientenfolge* $\{c_n\}$ *mit konvergentem* $\sum_{n=1}^{\infty} c_n^2 w_n$ *und fast überall divergentem* $\sum_{n=1}^{\infty} c_n f_n(x)$.[4])

Der folgende Satz bezieht sich auf $|C, \alpha|$-Summierbarkeit. Man sagt, die Reihe

$$\sum_{n=1}^{\infty} c_n \varphi_n(x) \qquad (\Sigma c_n^2 < \infty) \tag{2}$$

sei an der Stelle x $|C, \alpha|$-summierbar, wenn

$$\sum_{n=1}^{\infty} |\sigma_{n+1}^{(\alpha)}(x) - \sigma_n^{(\alpha)}(x)| < \infty$$

gilt, wobei $\sigma_n^{(\alpha)}(x)$ das n-te (C, α)-Mittel der Reihe (2) bezeichnet $(\alpha > -1)$.

Satz IV. *Damit die Haarsche Reihe* (1) *mit einer im Absolutbetrag monoton nichtwachsenden Koeffizientenfolge* $\{a_n\}$ *im Intervall* $(0, 1)$ *fast überall* $|C, \alpha|$*-summierbar ist, ist im Falle* $\alpha \geqq 0$ *die Bedingung*

$$\sum_{n=1}^{\infty} A_n < \infty, \quad A_n = \left(\sum_{k=2^n+1}^{2^{n+1}} a_k^2\right)^{1/2} \tag{3}$$

und im Falle $-1 < \alpha < 0$ *die Bedingung*

$$\sum_{n=1}^{\infty} 2^{-n\alpha} A_n < \infty \tag{4}$$

notwendig und hinreichend.

Tandori [6] hat bewiesen, daß die Bedingung (3) notwendig und hinreichend ist, damit die Reihe (2) für *jedes* orthonormierte Funktionensystem $|C, 1|$-summierbar ist. Danach hat der Verfasser [4] bewiesen, daß die Bedingung (3) bzw. (4) auch ohne Monotonie von $\{|a_n|\}$ die $|C, \alpha \geqq 0|$ bzw. $|C, \alpha \leqq 0|$-Summierbarkeit für jedes System vom Haarschen Typ sichert. P. Billard [2] hat nach der Erscheinung der Arbeit von Tandori bewiesen, daß die Rademachersche Reihe $\Sigma a_n r_n(x)$ auch dann nicht fast überall $|C, 1|$-summierbar ist, wenn die Bedingung (3) nicht erfüllt ist.

F. Móricz fragt in seiner Arbeit [5], ob die Bedingung (3) für eine positive, monoton nichtwachsende Koeffizientenfolge $\{a_n\}$ notwendig dafür ist, daß die Haarsche Reihe (1) (welche „besser" als die Rademachersche Reihe ist) fast überall $|C, 1|$-summierbar ist?

Der Satz IV gibt u. a. eine bejahende Antwort auf diese Frage.

Ferner sieht man, daß gewisse Sätze über Systeme vom Haarschen Typ, laut deren eine Bedingung für das Erfülltsein einer gewissen Konvergenz- bzw. Summationseigenschaft für die Gesamtheit der Systeme vom Haarschen Typ notwendig ist, für das Haarsche System selber gelten, wenn nur die Folge $\{|a_n|\}$ monoton ist. Und es gilt auch umgekehrt: Sätze, die behaupten, daß eine Bedingung unter der

---

[3]) $L_n(x)$ bezeichnet die n-te Lebesguesche Funktion von $\{f_n(x)\}$.

[4]) Im Satz von Tandori wird noch die Bedingung $\lambda_n \to \infty$ gestellt.

Monotonität der Folge $\{|a_n|\}$ für das Erfülltsein einer gewissen Konvergenz- bzw. Summationseigenschaft des Haarschen Systems notwendig ist, gelten ohne die Monotonität von $\{|a_n|\}$ nur für die Gesamtheit der Systeme vom Haarschen Typ.

§ 1. Beweis von Satz I

Wir nehmen an, daß die Reihe (1) mit einer monoton nichtwachsenden Folge $\{|a_n|\}$ auf einer Menge von positivem Maß konvergiert. Das bedeutet nach dem Egoroffschen Satz, daß es eine Menge E vom Maß $\mu(E)>0$ gibt, auf welcher die Partialsummen $s_n(x)$ die Eigenschaft

$$|s_{2^m}(x)-s_{2^n}(x)| \leqq K_1 \qquad (m>n)$$

haben, wo K_1 eine geeignete absolute Konstante ist. Daraus folgt

$$\mu(E)K_1^2 \geqq \int_E \left(\sum_{k=2^n+1}^{2^m} a_k\chi_k(x)\right)^2 dx \geqq \tag{1.1}$$

$$\geqq \mu(E)\sum_{k=n}^{m-1} 2^k a_{2^k+1}^2 - 2\left|\sum_{k=2^n+1}^{2^m-1} a_k \sum_{l=k+1}^{2^m} a_l \int_E \chi_k(x)\chi_l(x)\,dx\right|.$$

Bezeichne $I(k,p)$ die Menge der ganzen Zahlen l mit $k<l\leqq p$ und mit $\chi_k(x)\chi_l(x)\neq 0$ auf einer Menge von positivem Maß. Es ist klar, daß die letzte Doppelsumme in (1. 1) gleich

$$\left|\sum_{k=2^n+1}^{2^m-1} a_k \sum_{l\in I(k,2^m)} a_l \int_E \chi_k(x)\chi_l(x)\,dx\right|$$

ist.

Diese Summe ist nach der Cauchyschen Ungleichung nicht größer als[5])

$$\left\{\sum_{k=2^n+1}^{2^m-1} a_k^2 2^{[\log k]} \sum_{l\in I(k,2^m)} a_l^2\right\}^{1/2} \cdot \left\{\sum_{k=2^n+1}^{\infty} \frac{1}{2^{[\log k]}} \sum_{l\in I(k,\infty)} \left[\int_E \chi_k(x)\chi_l(x)\,dx\right]^2\right\}^{1/2}. \tag{1.2}$$

Man kann leicht einsehen, daß das System $\{2^{-[\log k]/2}\chi_k(x)\chi_l(x)\}$ mit $l\in I(k,\infty)$ in $(0, 1)$ orthonormiert ist. Also ist

$$2^{-[\log k]/2}\int_E \chi_k(x)\chi_l(x)\,dx$$

ein Entwicklungskoeffizient der charakteristischen Funktion $E(x)$ der Menge E. Nach der Besselschen Ungleichung folgt

$$\sum_{k=1}^{\infty}\sum_{l\in I(k,\infty)} 2^{-[\log k]}\left[\int_E \chi_k(x)\chi_l(x)\,dx\right]^2 \leqq \int_0^1 E^2(x)\,dx = \mu(E).$$

---

[5]) In dieser Arbeit benützen wir den Logarithmus mit der Basis 2.

Somit gilt für genügend große n die Ungleichung

$$(1.3)\qquad \sum_{k=2^n+1}^{\infty} 2^{-[\log k]} \sum_{l\in I(k,\infty)} \left[\int_E \chi_k(x)\chi_l(x)\,dx\right]^2 \leqq \frac{\mu(E)^2}{64}.$$

Ferner gilt die Abschätzung

$$(1.4)\qquad \sum_{k=2^n+1}^{2^m-1} a_k^2 2^{[\log k]} \sum_{l\in I(k,2^m)} a_l^2 \leqq \sum_{\nu=n}^{m-2} 2^\nu a_{2^\nu}^2 \sum_{\mu=\nu+1}^{m-1} 2^\mu a_{2^\mu}^2 \leqq$$

$$\leqq \sum_{\nu=n}^{m-2} 2^\nu a_{2^\nu}^2 \sum_{\mu=n+1}^{m-1} 2^\mu a_{2^\mu}^2 \leqq \left(\sum_{\nu=n}^{m-1} 2^\nu a_{2^\nu}^2\right)^2.$$

Nach (1. 2), (1. 3) und (1. 4) ergibt sich für genügend große n

$$2\left|\sum_{k=2^n+1}^{2^m-1} a_k \sum_{l=k+1}^{2^m} a_l \int_E \chi_k(x)\chi_l(x)\,dx\right| \leqq \frac{\mu(E)}{4}\sum_{\nu=n}^{m-1} 2^\nu a_{2^\nu}^2.$$

Aus (1. 1) erhält man dann

$$\mu(E)K_1^2 \geqq \mu(E)\sum_{k=n}^{m-1} 2^k a_{2^{k+1}}^2 - \frac{\mu(E)}{2}\sum_{k=n-1}^{m-2} 2^k a_{2^{k+1}}^2 \geqq$$

$$\geqq \frac{\mu(E)}{2}\sum_{k=n}^{m-1} 2^k a_{2^{k+1}}^2 - \mu(E)2^{n-2}a_{2^n}^2.$$

Da hier m beliebig ist, folgt somit

$$\sum_{k=n}^{\infty} 2^k a_{2^{k+1}}^2 = O(1),$$

folglich ist die Reihe $\sum_{k=1}^{\infty} a_k^2$ konvergiert.

Damit ist der Satz I bewiesen.

§ 2. Beweis von Satz II

Zuerst definieren wir ein spezielles orthonormiertes System vom Haarschen Typ $\{h_n(x)\}$ $(n=1, 2, \ldots)$. Seien $h_n(x) = \operatorname{sign}\sin 2^n\pi x$ $(n=1, 2)$. Es sei s $(\geqq 1)$ eine beliebige natürliche Zahl. Wir nehmen an, daß die Treppenfunktionen $h_n(x)$ $(n=1, 2, \ldots, 2^s)$ schon definiert sind, derart, daß sie ein System vom Haarschen Typ bilden, ferner mit $z(k, l; m, p; x) = h_{2^k+l}(x)h_{2^m+p}(x)$

$$(2.1)\qquad \int_0^1 z^2(k, l; m, p; x)\,dx = 1 \qquad (k<m\leqq s)$$

und für $(k, l, m, p) \not\equiv (k', l', m', p')$

$$\int_0^1 z(k, l; m, p; x) z(k', l'; m', p'; x)\, dx = 0 \tag{2.2}$$

gelten.

Dann kann das Intervall (0, 1) in endlich viele Teilintervalle $J_\varrho(s)$ $(1 \leqq \varrho \leqq \varrho_s)$ derart zerlegt werden, daß in jedem $J_\varrho(s)$ die Funktionen $h_n(x)$ $(n = 1, 2, \ldots, 2^s)$ konstant sind. Ohne Beschränkung der Allgemeinheit können wir annehmen, daß a_n für jedes n von 0 verschieden ist.

Wir setzen

$$\varrho_0(m) = 0 \quad \text{und} \quad \varrho_k(m) = \sum_{n=1}^{k} \frac{a_{2^m+n}^2}{A_m^2} \qquad (k = 1, 2, \ldots, 2^m;\ m = 1, 2, \ldots).$$

Für ein endliches Intervall $I = (u, v)$ wird

$$I_k(m; I) = \big(u + \mu(I)\varrho_{k-1}(m),\ u + \mu(I)\varrho_k(m)\big) \qquad (k = 1, 2, \ldots, 2^m;\ m = 1, 2, \ldots)$$

gesetzt.

Wir teilen jedes $J_\varrho(s) = (u_\varrho, v_\varrho)$ in 2^s Teilintervalle ein:

$$I_k(s; J_\varrho(s)) = \big(u_\varrho + \mu(J_\varrho(s))\varrho_{k-1}(s),\ u_\varrho + \mu(J_\varrho(s))\varrho_k(s)\big) \qquad (k = 1, 2, \ldots, 2^s)$$

mit

$$\mu\big(I_k(s; J_\varrho(s))\big) = \frac{a_{2^s+k}^2}{A_s^2}\,\mu\big(J_\varrho(s)\big),$$

und setzen

$$h_{2^s+k}(x) = \frac{A_s}{a_{2^s+k}} \sum_{\varrho=1}^{\varrho_s} h_1\big(x; I_k(s; J_\varrho(s))\big)\,^{6)} \qquad (k = 1, 2, \ldots, 2^s).$$

Die Funktionen $h_n(x)$ $(2^s < n \leqq 2^{s+1})$ sind Treppenfunktionen. Sie sind normiert:

$$\int_0^1 h_{2^s+k}^2(x)\, dx = \frac{A_s^2}{a_{2^s+k}^2} \sum_{\varrho=1}^{\varrho_s} \int_0^1 h_1^2\big(x; I_k(s; J_\varrho(s))\big)\, dx =$$

$$= \frac{A_s^2}{a_{2^s+k}^2} \sum_{\varrho=1}^{\varrho_s} \mu\big(I_k(s; J_\varrho(s))\big) \int_0^1 h_1^2(x)\, dx = \frac{A_s^2}{a_{2^s+k}^2} \sum_{\varrho=1}^{\varrho_s} \mu\big(J_\varrho(s)\big) \frac{a_{2^s+k}^2}{A_s^2} = 1.$$

---

[6]) Ist $I = (u, v)$ ein endliches Intervall und $h(x)$ eine in (0, 1) definierte Funktion, so wird

$$h(x; I) = \begin{cases} h\left(\dfrac{x-u}{v-u}\right) & \text{für } u < x < v, \\ 0 & \text{sonst} \end{cases}$$

gesetzt. Offensichtlich ist

$$\int_u^v h(x; I)\, dx = \mu(I) \int_0^1 h(x)\, dx.$$

Nach der Definition ist es klar, daß das Funktionensystem $\{h_n(x)\}$ $(n=1, 2, \dots, 2^{s+1})$ auch orthogonal ist und für jedes $x \in (0, 1)$ sogar

$$h_n(x) h_m(x) = 0 \qquad (2^l < n, m \leqq 2^{l+1}, n \neq m;\ 0 \leqq l \leqq s)$$

ist, also bilden die $h_n(x)$ $(n=1, 2, \dots, 2^{s+1})$ ein System von Haarschem Typ.

Durch eine einfache Rechnung kann man einsehen, daß (2. 1) auch mit $s+1$ gilt. Sei $k < m \leqq s+1$, dann ist

$$\int_0^1 z^2(k, l; m, p; x)\, dx =$$

$$= \frac{A_k^2}{a_{2^k+l}^2} \cdot \frac{A_m^2}{a_{2^m+p}^2} \sum_{\varrho=1}^{\varrho_k} \int\limits_{I_l(k;\, J_\varrho(k))} \sum_{\varrho'=1}^{\varrho_m} h_1^2(x; I_p(m; J_{\varrho'}(m))\, dx =$$

$$= \frac{A_k^2}{a_{2^k+1}^2} \cdot \frac{A_m^2}{a_{2^m+p}^2} \sum_{\varrho=1}^{\varrho_k} \mu(I_l(k; J_\varrho(k))) \frac{a_{2^m+p}^2}{A_m^2} = \frac{A_k^2}{a_{2^k+l}^2} \sum_{\varrho=1}^{\varrho_k} \frac{a_{2^k+l}^2}{A_k^2} \mu(J_\varrho(k)) = 1.$$

Das Erfülltsein von (2. 2) ist nach der Definition der Funktionen $h_n(x)$ klar.

Durch vollständige Induktion ergibt sich sodann ein unendliches System von Haarschem Typ mit den Eigenschaften (2. 1) und (2. 2).

Der Beweis verläuft analog zu dem Beweis des Satzes I. Wir nehmen an, daß die Reihe

$$\sum_{n=1}^{\infty} a_n h_n(x)$$

auf einer Menge von positivem Maß konvergiert. Nach dem Egoroffschen Satz gibt es dann eine Menge E vom Maß $\mu(E) > 0$, auf welcher die Partialsummen $s_n(x)$ die Eigenschaft

$$|s_{2^m}(x) - s_{2^n}(x)| \leqq K_2 \qquad (m > n)$$

haben, wobei K_2 eine Konstante ist. Deswegen besteht nach der Definition der Funktionen $h_n(x)$

$$\mu(E) K_2^2 \geqq \int_E \Bigl(\sum_{k=2^n+1}^{2^m} a_k h_k(x) \Bigr)^2 dx \geqq \tag{2. 3}$$

$$\geqq \mu(E) \sum_{k=n}^{m-1} A_k^2 - 2 \left| \sum_{k=2^n+1}^{2^m-1} a_k \sum_{l=k+1}^{2^m} a_l \int_E h_k(x) h_l(x)\, dx \right|.$$

Da das System vom Haarschen Typ ist, ist die rechts stehende Doppelsumme mit

$$\left| \sum_{k=2^n+1}^{2^m-1} a_k \sum_{l=2^{[\log(k-1)]+1}+1}^{2^m} a_l \int_E h_k(x) h_l(x)\, dx \right|$$

gleich. Diese Summe kann man nach der Cauchyschen Ungleichung folgenderweise abschätzen:

$$\text{(2. 4)} \qquad \left\{\sum_{k=2^n+1}^{2^m-1} a_k^2 \sum_{l=2^{[\log(k-1)]+1}+1}^{2^m} a_l^2\right\}^{1/2} \cdot$$

$$\cdot\left\{\sum_{k=2^n+1}^{2^m-1} \sum_{l=2^{[\log(k-1)]+1}+1}^{2^m} \left[\int_E h_k(x) h_l(x)\,dx\right]^2\right\}^{1/2}.$$

Nach (2. 1) und (2. 2) ist das System $\{h_k(x) h_l(x)\}$ $l > 2^{[\log(k-1)]+1}$ $k = 1, 2, \ldots$ orthonormiert. Das Integral

$$\int_E h_k(x) h_l(x)\,dx$$

ist ein Entwicklungskoeffizient der charakteristischen Funktion $E(x)$ der Menge E. Nach der Besselschen Ungleichung ergibt sich

$$\sum_{k=2}^{\infty} \sum_{l=2^{[\log(k-1)]+1}+1}^{\infty} \left[\int_E h_k(x) h_l(x)\,dx\right]^2 \leqq \int_0^1 E^2(x)\,dx = \mu(E).$$

Wenn n genügend groß ist, besteht

$$\text{(2. 5)} \qquad \sum_{k=2^n+1}^{\infty} \sum_{l=2^{[\log(k-1)]+1}+1}^{\infty} \left[\int_E h_k(x) h_l(x)\,dx\right]^2 \leqq \frac{\mu(E)^2}{16}.$$

Nach (2. 4) und (2. 5) gilt,

$$2\left|\sum_{k=2^n+1}^{2^m-1} a_k \sum_{l=2^{[\log(k-1)]+1}+1}^{2^m} a_l \int_E h_k(x) h_l(x)\,dx\right| \leqq \frac{\mu(E)}{2} \sum_{k=2^n+1}^{2^m} a_k^2,$$

somit ergibt sich aus (2. 3)

$$\mu(E) K_2^2 \geqq \mu(E) \sum_{k=n}^{m-1} A_k^2 - \frac{\mu(E)}{2} \sum_{k=n}^{m-1} A_k^2 = \frac{\mu(E)}{2} \sum_{k=n}^{m-1} A_k^2.$$

Da m beliebig groß sein kann, so konvergiert $\sum_{k=1}^{\infty} a_k^2$.

Damit haben wir den Satz II bewiesen.

§ 3. Beweis von Satz III

Den Satz III brauchen wir jetzt nur im Falle $\lambda_n = O(1)$ zu beweisen. Ist $\lambda_n = O(1)$, so folgt aus $w_n = o(\lambda_n)$, daß die Folge $\{w_n\}$ gegen Null strebt. Es ist klar, daß eine monoton abnehmende Koeffizientenfolge $\{c_n\}$ mit $\Sigma c_n^2 = \infty$ und $\Sigma c_n^2 w_n < \infty$ angegeben werden kann. Es sei $f_n(x)=\chi_n(I; x)$ mit $I=(a, b)$. Die Reihe $\Sigma c_n f_n(x)$ divergiert nach dem Satz I fast überall. Ferner besteht

$$\int_a^b \left| \sum_{n=0}^{m} \chi_n(I; x)\chi_n(I; t) \right| dt \leqq \mu(I) \qquad (m = 0, 1, \ldots)$$

nach dem bekannten Ergebnis von A. HAAR [3], also sind die Lebesgueschen Funktionen von $\{f_n(x)\}$ beschränkt.

Damit ist der Satz III vollständig bewiesen.

§ 4. Beweis von Satz IV

Die Hinlänglichkeit der Bedingungen (3) und (4) folgt aus Satz V in [4]. Zum Beweis der Notwendigkeit benötigen wir den folgenden Hilfssatz aus [4]:

Hilfssatz I. *Es sei $\{R_n(x)\}_1^\infty$ ein im Intervall $(0, 1)$ definiertes Treppenfunktionensystem. Wir bezeichnen mit $J_s(n)$ $(n=1, 2, \ldots; s=1, 2, \ldots, s_n)$ die Konstanzintervalle von $R_n(x)$. Gilt für jedes $m>n$ die Beziehung*

$$\int_{J_s(n)} \operatorname{sign} R_m(x)\, dx = 0 \qquad (s=1, 2, \ldots, s_n),$$

so gibt es zu jeder reellen Zahlenfolge $d_1, \ldots, d_N$ einfache Mengen[7]) *E_k $(k=0, 1, \ldots, N-1)$ derart, daß für jedes $x \in E_k$*

$$\left| \sum_{l=1}^{N} d_l R_l(x) \right| \geqq |d_{N-k} R_{N-k}(x)|$$

besteht und

$$\mu(E_k \cap J_s(N-k-1)) = \frac{\mu(J_s(N-k-1))}{2^{k+1}} \text{ [8]}) \qquad (s=1, 2, \ldots, s_{N-k-1};\ J_1(0)=(0, 1))$$

gilt.

Bekanntlich ist $c_1(\alpha) \leqq \dfrac{A_m^{(\alpha)}}{m^\alpha} \leqq c_2(\alpha)$ $(m>0, \alpha>-1)$, wobei $c_1(\alpha)$ und $c_2(\alpha)$ nur von α abhängige positive Zahlen sind, ferner gilt für

$$L_{n,\nu}^{(\alpha)} = \frac{A_{n+1-\nu}^{(\alpha)}}{A_{n+1}^{(\alpha)}} - \frac{A_{n-\nu}^{(\alpha)}}{A_n^{(\alpha)}} = \frac{A_{n-\nu}^{(\alpha)}}{A_n^{(\alpha)}} \frac{\nu\alpha}{(n+1-\nu)(n+1+\alpha)}$$

die Abschätzung

$$d_1(\alpha) \frac{(n+1-\nu)^{\alpha-1}\nu}{n^{\alpha+1}} \leqq |L_{n,\nu}^{(\alpha)}| \leqq d_2(\alpha) \frac{(n+1-\nu)^{\alpha-1}\nu}{n^{\alpha+1}} \tag{4. 1}$$

[7]) D. h. E_k ist die Summe endlich vieler Intervalle.

[8]) Mit $\mu(H)$ wird das Lebesguesche Maß der Menge H bezeichnet.

($n=1, 2, \ldots; \nu=0,1, \ldots, n; \alpha>-1, \alpha\neq 0$), wobei $d_1(\alpha)$ und $d_2(\alpha)$ nur von α abhängige positive Zahlen sind.

Wir nehmen an, daß die Reihe (1) fast überall $|C, \alpha|$-summierbar ist. Es sei $\varepsilon=2^{-(24+2\alpha)}d_1^2(\alpha)d_2^{-2}(\alpha)$, wo $d_1(\alpha)$ und $d_2(\alpha)$ die unter (4. 1) angeführten Konstanten sind. Nach dem Egoroffschen Satz gibt es dann eine meßbare Menge E mit $\mu(E)\geqq 1-\varepsilon$ und eine positive Konstante K derart, daß für die (C, α)-Mitteln der Reihe (1)

$$\sum_{n=1}^{\infty}|\sigma_{n+1}^{(\alpha)}(x)-\sigma_n^{(\alpha)}(x)|<K \qquad (x\in E)$$

ist. Hieraus folgt:

$$\sum_{n=2}^{\infty}\int_E|\sigma_{n+1}^{(\alpha)}(x)-\sigma_n^{(\alpha)}(x)|\,dx\leqq K\mu(E). \tag{4. 2}$$

Seien m und n natürliche Zahlen mit $2^m<n\leqq 2^{m+1}$. Wir setzen

$$R_l(x; m, n)=\sum_{\nu=2^l+1}^{2^{l+1}}L_{n,\nu}^{(\alpha)}a_\nu\chi_\nu(x) \qquad (l=1, 2, \ldots, m-1),$$

$$R_m(x; m, n)=\sum_{\nu=2^m+1}^{n}L_{n,\nu}^{(\alpha)}a_\nu\chi_\nu(x),$$

und

$$R_{m+1}(x; m, n)=\frac{1}{A_{n+1}^{(\alpha)}}a_{n+1}\chi_{n+1}(x).$$

Die Funktionen $R_l(x; m, n)$ ($l=1, 2, \ldots, m+1$) erfüllen offensichtlich die Bedingungen des Hilfssatzes I.

Wenn $\alpha\geqq 0$ ist, dann wenden wir auf die Funktionen $R_l(x; m, n)$ ($l=1, 2, \ldots, m+1$) den Hilfssatz I mit $N=m+1$, $k=3$ an; die entsprechende Menge sei mit $E_3=E_3(m, n)$ bezeichnet. So ergibt sich

$$\sum_{n=0}^{\infty}\int_E|\sigma_{n+1}^{(\alpha)}(x)-\sigma_n^{(\alpha)}(x)|\,dx=\sum_{m=3}^{\infty}\sum_{n=2^m+1}^{2^{m+1}}\int_E\Bigg|\sum_{\nu=0}^{n}L_{n,\nu}^{(\alpha)}a_\nu\chi_\nu(x)+$$

$$+\frac{1}{A_{n+1}^{(\alpha)}}a_{n+1}\chi_{n+1}(x)\Bigg|\,dx\geqq\sum_{m=3}^{\infty}\sum_{n=2^m+1}^{2^{m+1}}\int_{E_3\cap E}\Bigg|\sum_{\nu=2^{m-2}+1}^{2^{m-1}}L_{n,\nu}^{(\alpha)}a_\nu\chi_\nu(x)\Bigg|\,dx\geqq$$

$$\geqq\sum_{m=3}^{\infty}\sum_{n=2^m+1}^{2^{m+1}}\Bigg(\int_{E_3}-\int_{E_3-E_3\cap E}\Bigg)\Bigg|\sum_{\nu=2^{m-2}+1}^{2^{m-1}}L_{n,\nu}^{(\alpha)}a_\nu\chi_\nu(x)\Bigg|\,dx=\Sigma_1.$$

Mit Anwendung der Schwarzschen Ungleichung bekommt man nach (4. 1)

$$\Sigma_1 \geqq \sum_{m=3}^{\infty} \sum_{n=2^m+1}^{2^{m+1}} \left(\int_{E_3} \left| \sum_{\nu=2^{m-2}+1}^{2^{m-1}} L_{n,\nu}^{(\alpha)} a_\nu \chi_\nu(x) \right| dx - \right.$$

$$\left. - \sqrt{\varepsilon} \left\{ \int_0^1 \left| \sum_{\nu=2^{m-2}+1}^{2^{m-1}} L_{n,\nu}^{(\alpha)} a_\nu \chi_\nu(x) \right|^2 dx \right\}^{1/2} \right) \geqq$$

$$\geqq \sum_{m=3}^{\infty} \sum_{n=2^m+1}^{2^{m+1}} \left(2^{-4} \sum_{\nu=2^{m-2}+1}^{2^{m-1}} L_{n,\nu}^{(\alpha)} |a_\nu| \frac{1}{2^{m/2}} - \sqrt{\varepsilon}\, d_2(\alpha) 2^{-(\alpha+m)} A_{m-2} \right) \geqq$$

$$\geqq \sum_{m=3}^{\infty} \left(2^{-7} d_1(\alpha) \frac{1}{2^{m/2}} \sum_{\nu=2^{m-2}+1}^{2^{m-1}} |a_\nu| \sum_{n=2^m+1}^{2^{m+1}} \frac{(n+1-\nu)^{\alpha-1}}{n^\alpha} - \right.$$

$$\left. - \sqrt{\varepsilon} d_2(\alpha) 2^{-\alpha} A_{m-2} \right) = \Sigma_2 .$$

Nach der Definition von ε besteht

$$\Sigma_2 \geqq \sum_{m=3}^{\infty} \left(2^{-(8+2\alpha)} d_1(\alpha) 2^{-\frac{m}{2}} \sum_{\nu=2^{m-2}+1}^{2^{m-1}} |a_\nu| - 2^{-(11+2\alpha)} d_1(\alpha) 2^{\frac{m}{2}} |a_{2^{m-2}}| \right) \geqq$$

$$\geqq \sum_{m=3}^{\infty} (2^{-(10+2\alpha)} d_1(\alpha) 2^{m/2} |a_{2^{m-1}}| - 2^{-(12+2\alpha)} d_1(\alpha) 2^{m/2} |a_{2^{m-2}}|) \geqq$$

$$\geqq \sum_{m=3}^{\infty} 2^{-(11+2\alpha)} d_1(\alpha) 2^{m/2} |a_{2^{m-1}}| - 2^{-(11+2\alpha)} d_1(\alpha) 2^{3/2} |a_2| \geqq$$

$$\geqq \sum_{m=2}^{\infty} 2^{-(11+2\alpha)} d_1(\alpha) A_m - O(1).$$

Nach obigen folgt das Erfülltsein der Bedingung (3) wegen (4. 2).

Es sei $-1<\alpha<0$. Bei Benützung der vorigen Bezeichnungen nehmen wir an, daß die Reihe (1) in (0, 1) fast überall $|C, \alpha|$-summierbar ist. Dann gibt es auch jetzt nach dem Egoroffschen Satz zu $\varepsilon = 2^{-12} d_1^2(\alpha) d_2^{-2}(\alpha)$ eine meßbare Menge E mit $\mu(E) \geqq 1-\varepsilon$ und eine Konstante M derart, daß

$$\sum_{n=0}^{\infty} |\sigma_{n+1}^{(\alpha)}(x) - \sigma_n^{(\alpha)}(x)| < M \qquad (x \in E)$$

ist, woraus

$$\sum_{n=0}^{\infty} \int_E |\sigma_{n+1}^{(\alpha)}(x) - \sigma_n^{(\alpha)}(x)|\, dx \leqq M\mu(E) \tag{1. 3}$$

folgt. Wir wenden den Hilfssatz I für die $R_l(x; m, n)$ mit $N = m+1$ und $k=1$ an; die entsprechende Menge bezeichnen wir mit $E_1 = E_1(m, n)$.

Ähnlich wie vorher ergibt sich

$$\sum_{n=4}^{\infty} \int_E |\sigma_{n+1}^{(\alpha)}(x) - \sigma_n^{(\alpha)}(x)|\, dx =$$

$$= \sum_{m=1}^{\infty} \sum_{n=2^m+1}^{2^{m+1}} \int_E \left| \sum_{\nu=0}^{n} L_{n,\nu}^{(\alpha)} a_\nu \chi_\nu(x) + \frac{1}{A_{n+1}^{(\alpha)}} a_{n+1}\chi_{n+1}(x) \right| dx \geqq$$

$$\geqq \sum_{m=1}^{\infty} \sum_{n=2^m+1}^{2^{m+1}} \int_{E_1 \cap E} \left| \sum_{\nu=2^m+1}^{n} L_{n,\nu}^{(\alpha)} a_\nu \chi_\nu(x) \right| dx \geqq$$

$$\geqq \sum_{m=1}^{\infty} \sum_{n=2^m+1}^{2^{m+1}} \left(\int_{E_1} - \int_{E_1 - E_1 \cap E} \right) \left| \sum_{\nu=2^m+1}^{n} L_{n,\nu}^{(\alpha)} a_\nu \chi_\nu(x) \right| dx = \Sigma'.$$

Die Abschätzung kann man durch Anwendung der Schwarzschen Ungleichung fortsetzen:

$$\Sigma' \geqq \sum_{m=1}^{\infty} \sum_{n=2^m+1}^{2^{m+1}} \left(\int_{E_1} \left| \sum_{\nu=2^m+1}^{n} L_{n,\nu}^{(\alpha)} a_\nu \chi_\nu(x) \right| dx - \right.$$

$$\left. - \sqrt{\varepsilon} \left\{ \int_0^1 \left| \sum_{\nu=2^m+1}^{n} L_{n,\nu}^{(\alpha)} a_\nu \chi_\nu(x) \right|^2 dx \right\}^{1/2} \right) \geqq$$

$$\geqq \sum_{m=1}^{\infty} \sum_{n=2^m+1}^{2^{m+1}} \left(2^{-2} \sum_{\nu=2^m+1}^{n} |L_{n,\nu}^{(\alpha)}|\, |a_\nu| 2^{-\frac{m}{2}} - \sqrt{\varepsilon}\, d_2(\alpha) 2^{-\alpha m} A_m \right) \geqq$$

$$\geqq \sum_{m=1}^{\infty} \left(2^{-3} d_1(\alpha) 2^{-\frac{m}{2}} \sum_{\nu=2^m+1}^{2^{m+1}} |a_\nu| \sum_{n=\nu}^{2^{m+1}} \frac{(n+1-\nu)^{\alpha-1}}{n^\alpha} - \sqrt{\varepsilon}\, d_2(\alpha) 2^{-\alpha m} A_m \right) \geqq$$

$$\geqq \sum_{m=1}^{\infty} \left(2^{-3} d_1(\alpha) 2^{m\left(\frac{1}{2}-\alpha\right)} |a_{2^{m+1}}| - \sqrt{\varepsilon}\, d_2(\alpha) 2^{m\left(\frac{1}{2}-\alpha\right)} |a_{2^m}| \right) = \Sigma''.$$

Nach der Definition von ε gilt

$$\Sigma'' \geqq -2^{-4} d_1(\alpha) |a_2| + \sum_{m=2}^{\infty} \left(2^{-6} d_1(\alpha) 2^{m\left(\frac{1}{2}-\alpha\right)} |a_{2^m}| \right) \geqq -O(1) + O(1) \sum_{m=2}^{\infty} 2^{-m\alpha} A_m,$$

woraus die Bedingung (4) auf Grund von (4. 3) folgt.

Damit haben wir den Satz IV vollständig bewiesen.

Schriftenverzeichnis

[1] G. ALEXITS, *Konvergenzprobleme der Orthogonalreihen* (Budapest, 1960).
[2] P. BILLARD, Sur la sommabilité absolue des séries de fonctions orthogonales, *Bull. Sci. Math.*, **85** (1961), 29–33.
[3] A. HAAR, Zur Theorie der orthogonalen Funktionensysteme, *Math. Annalen*, **69** (1910), 331–371.
[4] L. LEINDLER, Über die absolute Summierbarkeit der Orthogonalreihen, *Acta Sci. Math.*, **22** (1961), 243–268.
[5] Ф. Мориц, О безусловной сходимости рядов по системе Хаара, *Известия Акад. Наук СССР*, **27** (1963), 1229–1238.
[6] K. TANDORI, Über die orthogonalen Funktionen. IX (Absolute Summation), *Acta Sci. Math.*, **21** (1960), 292–299.
[7] K. TANDORI, Über die orthogonalen Funktionen. V (Genaue Weylsche Multiplikatorfolgen), *Acta Sci. Math.*, **20** (1959), 1–13.

(Eingegangen am 23. April 1964)

On completely continuous and uniformly bounded operators in l^p spaces

By L. GEHÉR in Szeged

To Professor L. Kalmár on his 60th birthday

Introduction and terminology. In the sequel, a *transformation* T means a linear, bounded (and everywhere defined) mapping of a Banach space $\mathfrak{B}_1$ into another Banach space $\mathfrak{B}_2$; T is called *invertible* if it has a bounded (everywhere defined) inverse T^{-1}. In case $\mathfrak{B}_1 = \mathfrak{B}_2$, T is said to be an *operator*. An operator T defined on a Banach space is called uniformly bounded if there exists a constant K with $\|T^n\| \leqq K$ $(n=0, 1, \ldots)$. *Contraction* means an operator with norm not greater than 1. Two operators A and B defined on $\mathfrak{A}$ and $\mathfrak{B}$, respectively, are said to be *similar* if there exists an invertible transformation C mapping $\mathfrak{A}$ onto $\mathfrak{B}$ such that $A = C^{-1}BC$.

A few years ago B. Sz.-Nagy [1] proved the following theorem:

Every completely continuous and uniformly bounded operator defined on a Hilbert space $\mathfrak{H}$ is similar to a contraction of $\mathfrak{H}$.

The theorem also has been proved for l_n^p spaces [2].

In this paper we are going to give a generalization of the above mentioned theorem of B. Sz.-Nagy in l^p spaces. The following theorem will be proved:

Theorem. *Every completely continuous and uniformly bounded operator on l^p $(1 \leqq p < \infty)$ is similar to a contraction defined on a proper subspace of l^p isomorphic to l^p.*

Every subspace of a Hilbert space $\mathfrak{H}$ isomorphic to $\mathfrak{H}$ being also isometric to $\mathfrak{H}$, our result contains that of B. Sz.-Nagy (at least for separable Hilbert spaces).

Proof of the theorem. 1°. Let T be a completely continuous and uniformly bounded operator defined on l^p. Then the spectrum Σ of T consists of countable many points, the only possible accumulation point of which is the point 0. Further, T being uniformly bounded, its spectrum Σ lies on the closed unit disc, and on the unit circle there are only finitely many eigenvalues of T. Denote by $\mu_1, \mu_2, \ldots, \mu_k$ the eigenvalues lying on the unit circle, each of them being written as many times as its geometric multiplicity (for a uniformly bounded operator, each eigenvalue lying on the unit circle has the same algebraic multiplicity as its geometric multiplicity, cf. [1]). Let $a<1$ be a positive number such that the rest Σ' of the spectrum lies in the disc $|z|<a$. Then, as well-known,

$$P=-\frac{1}{2\pi i}\oint_{|z|=a} R_z\,dz, \quad \text{where} \quad R_z=(T-zI)^{-1},$$

is a projection commuting with T, such that the spectrum of the restriction $T_{\mathfrak{B}}$ of T to $\mathfrak{B}=Pl^p$ is Σ'. $\mathfrak{B}$ has finite codimension. (A subspace of a Banach space has finite codimension k, if it has a k-dimensional complementary subspace.) $\mathfrak{B}$ being isomorphic to l^p, there exists an operator T' on l^p similar to $T_{\mathfrak{B}}$.*) For a moment, suppose that the theorem is true for T', i. e. that there exists a subspace $\mathfrak{B}_1'$ of l^p, such that T' is similar to a contraction on $\mathfrak{B}_1'$.

The space of the elements $(\alpha_1,\alpha_2,\ldots,\alpha_k,x)$, where $\alpha_1,\alpha_2,\ldots,\alpha_k$ are scalars and $x\in l^p$, supplied with the norm $\left(\sum_{i=1}^{k}|\alpha_i|^p+\|x\|^p\right)^{1/p}$ is a Banach space isometric to l^p, thus it can be identified with l^p. The elements $(\alpha_1,\alpha_2,\ldots,\alpha_k,x)$ $(x\in\mathfrak{B}_1')$ form a subspace $\mathfrak{B}_1$ of this space. The operator

$$T_1(\alpha_1,\alpha_2,\ldots,\alpha_k,x)=(\mu_1\alpha,\mu_2\alpha_2,\ldots,\mu_k\alpha_k,T_1'x) \qquad (x\in\mathfrak{B}_1')$$

defined on this subspace is evidently a contraction and similar to T.

---

*) We have made use of the fact that every finite codimensional subspace of l^p is isomorphic to l^p. Although this is a consequence of a general theorem of Pełczyński [3], in this special case we sketch a proof for it. Let $\mathfrak{B}$ be a k-codimensional subspace of l^p (k is finite, $k>0$) and P a (bounded) projection of l^p onto $\mathfrak{B}$. Introducing the basis vectors $e_1=(1,0,0,\ldots)$, $e_2=(0,1,0,\ldots),\ldots$ in l^p, every element $x\in\mathfrak{B}$ may be written in the form

$$x=\sum_{n=1}^{\infty}\alpha_n Pe_n \quad \text{with} \quad \sum_{n=1}^{\infty}|\alpha_n|^p<\infty,$$

but not uniquely. It can be proved, however, that we can choose a set of k subscripts $N=\{n_1,n_2,\ldots,n_k\}$ such that a representation

$$x=\sum_{n\notin N}\alpha_n' Pe_n \quad \text{with} \quad \sum_{n\notin N}|\alpha_n'|^p<\infty$$

is possible and is uniquely determined, for each $x\in\mathfrak{B}$. I. e. $\{Pe_n\}_{n\notin N}$ is a basis for $\mathfrak{B}$ and P furnishes, by the closed graph theorem, an isomorphic mapping of the space spanned by the vectors e_n $(n\notin N)$ onto $\mathfrak{B}$. To construct such a set N consider a basis $q_i=\alpha_{i1}e_1+\alpha_{i2}e_2+\ldots$ $(i=1,2,\ldots,k)$ of the complementary subspace $(I-P)l^p$ of $\mathfrak{B}$. Then the matrix

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \cdots & \cdots \\ \alpha_{21} & \alpha_{22} & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{k1} & \alpha_{k2} & \cdots & \cdots & \cdots \end{pmatrix}$$

is of rank k, i. e. it has a subdeterminant of order k different from 0, e. g.

$$D=\begin{vmatrix} \alpha_{1s_1} & \alpha_{1s_2} & \cdots & \alpha_{1s_k} \\ \alpha_{2s_1} & \alpha_{2s_2} & \cdots & \alpha_{2s_k} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{ks_1} & \alpha_{ks_2} & \cdots & \alpha_{ks_k} \end{vmatrix}\neq 0.$$

An easy calculation with determinants shows that the set $N=\{s_1,s_2,\ldots,s_k\}$ satisfies the above requirement.

So it is enough to deal with the case when the spectrum is contained in the interior of the unit disc.

2°. Suppose that T has no eigenvalue lying on the unit circle. In virtue of the well-known theorem of Gelfand, $\lim_{n\to\infty} \|T^n\|^{1/n} = \varrho < 1$, where ϱ denotes the spectral radius of T. Thus for any number σ with $\varrho < \sigma < 1$ there exists a natural number N such that $\|T^n\| \leqq \sigma^n$ for $n > N$. We define a new norm in l^p with

$$|x| = \left(\sum_{n=0}^{\infty} \|T^n x\|^p\right)^{1/p}.$$

It follows from

$$\|x\| \leqq |x| = \left(\sum_{n=0}^{\infty} \|T^n x\|^p\right)^{1/p} \leqq \|x\| \left(\sum_{n=0}^{N} \|T^n\|^p + \sum_{n=N+1}^{\infty} \sigma^{np}\right)^{1/p} = C\|x\| \qquad (C < \infty)$$

that this new norm is equivalent to the old one. Denote by $\boldsymbol{l}^p$ this new space and by $\boldsymbol{T}$ the operator on $\boldsymbol{l}^p$ defined by $\boldsymbol{T}x = Tx$. Evidently, $\boldsymbol{T}$ is a contraction. The two norms being equivalent, $\boldsymbol{T}$ is similar to T. Now, we have only to prove that $\boldsymbol{l}^p$ may be isometrically imbedded into l^p. Let Θ be a rearrangement of the double sequence into a simple sequence. Rearranging the double sequence $(x, Tx, T^2x, \ldots)$, where $x, Tx, T^2x, \ldots$ stand instead of the sequences corresponding to $x, Tx, T^2x, \ldots$, respectively, by means of Θ, we easily get an isometric transformation of $\boldsymbol{l}^p$ into l^p.

Remark. Suppose that there exists a decreasing sequence $\mathfrak{B}_1 \supset \mathfrak{B}_2 \supset \ldots$ of finite codimensional subspaces of l^p reducing T, with $\bigcap_{n=1}^{\infty} \mathfrak{B}_n = 0$. (For instance, this is the case if the point 0 is not an eigenvalue of T. Indeed, let $\{a_n\}_1^\infty$ be a decreasing sequence of positive numbers, converging to 0, such that the circles with radius a_n $(n = 1, 2, \ldots)$ and centre 0 contain no eigenvalue of T. Then with the aid of the projections

$$P_n = -\frac{1}{2\pi i} \oint_{|z| = a_n} R_z \, dz$$

we get a sequence of invariant subspaces having the required property.) Then T is similar to a contraction defined on a *finite codimensional* subspace of l^p. The proof is based on the fact that if $T_1, T_2, \ldots$ denote the restrictions of T onto $\mathfrak{B}_1, \mathfrak{B}_2, \ldots$, respectively, then $\|T_n\| \to 0$ for $n \to \infty$. To prove this suppose the contrary, i.e. that there exist a positive number $\varepsilon > 0$ and a sequence $\{x_n\}_1^\infty$, with $\|x_n\| = 1$ and $x_n \in \mathfrak{B}_n$ $(n = 1, 2, \ldots)$ such that

$$\|T_n x_n\| = \|T x_n\| \geqq \varepsilon.$$

T being completely continuous, there exists a convergent subsequence $\{Tx_{i_n}\}_1^\infty$ of $\{Tx_n\}_1^\infty$. Let $\lim_{n\to\infty} Tx_{i_n} = x$. Then by $\|Tx_{i_n}\| \geqq \varepsilon$ it follows that $\|x\| \geqq \varepsilon$. On the other hand, in virtue of $x_i \in \mathfrak{B}_n$ and $\bigcap_{n=1}^{\infty} \mathfrak{B}_n = 0$, we get $x = 0$, which is a contradiction.

Using this, the proof can be finished so as in the former case.

References

B. Sz.-Nagy, Completely continuous operators with uniformly bounded iterates, *Publ. Math. Inst. Hung. Acad. Sci.*, **4** (1959), 89—93.
L. Gehér, Über ähnliche lineare Transformationen in endlichdimensionalen Räumen, *Publ. Math. Inst. Hung. Acad. Sci.*, **4** (1959), 95—99.
A. Pelczyński, Projections in certain Banach spaces, *Studia Math.*, **19** (1960), 209—228.

(Received January 31, 1965)

On partially ordered algebras. II

By L. FUCHS in Budapest

Dedicated to L. Kalmár on his 60th birthday

This note is the continuation of the author's paper "On partially ordered algebras. I" to appear in *Colloquium Mathematicum.* It will be shown that every algebra gives rise to a lattice-ordered algebra where the monotony domains coincide with the whole underlying set and all the operations are isotone in each of their variables.[1]) Some results on ideal systems can be generalized mutatis mutandis to our more general case.

1. Lattice-ordered algebras obtained from arbitrary algebras

We are going to show how from every algebra it is possible to derive a lattice-ordered algebra e. g. as the set of subalgebras or subsystems of a special kind. This is the way in which the lattice of subgroups, normal subgroups of groups, subrings, left ideals or ideals of rings, l-ideals of lattice-ordered groups, submodules of modules etc. are obtained.

In order to frame our definition more generally, we recall the concept of ideal systems. We formulate this here in a way which suites better to our purposes.

Let $(A; F)$ be an arbitrary algebra and

$$X \to X_r$$

a mapping from the set of all non-void finite subsets of A into the set of all subsets of A satisfying the following conditions:

1°. $X \subseteqq X_r$,

2°. $X \subseteqq Y_r$ implies $X_r \subseteqq Y_r$.

(That is to say, it satisfies the axioms of closure operators.) Usually, $(a)_r$ is prescribed in some way.

We extend the domain of definition of this r-operation to infinite subsets Z of A by setting [2])

3°. $Z_r = \bigcup X_r$ where X runs over all finite subsets of Z.

Then, as readily checked, axioms 1°—2° hold for infinite X, Y as well. The system

---

[1]) For the basic definitions on partially ordered algebras in general we refer to Part I. For partially ordered groups and rings see e. g. [1].

[2]) We shall use the signs $\wedge$, $\vee$ to denote lattice-operations, while $\cap$, $\cup$ will denote set-theoretical meet and join, respectively.

of all X_r with finite and infinite X will be denoted by A_r and called the *r-ideal system* of A.

A_r is partially ordered under inclusion. It is closed with respect to intersection — this is an immediate consequence of 1° and 2°. If we define

$$X_r \vee Y_r = (X \cup Y)_r$$

which is easily shown to be independent of the representation of X_r and Y_r by X and Y, respectively), then A_r becomes a lattice. In A_r we have evidently

$$\bigvee_\alpha X_r^\alpha = (\bigcup_\alpha X^\alpha)_r \tag{1}$$

with α running over an arbitrary index set.

For each $f \in F$ we set[3])

$$\bar{f}(X_r^1, \ldots, X_r^n) = (\cup f(x^1, \ldots, x^n) \quad \text{for all} \quad x^i \in X_r^i)_r;$$

then every $\bar{f}$ in A_r satisfies the monotony law with the whole of A_r as monotony domain. The operations $\bar{f}$ are evidently isotone in their variables. Hence A_r becomes a lattice-ordered algebra with the operations $\bar{f}$ $(f \in F)$ and $\wedge, \vee$; moreover A_r is a complete lattice under $\wedge, \vee$. However, not all the identities of A retain their validity in A_r.

Call an operation $f \in F$ *r-admissible* if for each i it satisfies

$$\begin{aligned}(f(x^1, \ldots, x^{i-1}, X_r^i, x^{i+1}, \ldots, x^n))_r = \\ = (f(x^1, \ldots, x^{i-1}, x^i, x^{i+1}, \ldots, x^n) \quad \text{for all} \quad x^i \in X^i)_r.\end{aligned} \tag{2}$$

Recall that in partially ordered groups or in commutative rings only r-ideal systems are considered which are supposed to have the property $(aX)_r = aX_r$ which is somewhat stronger than (2).

For r-admissible operations $f \in F$ we have:

$$\bar{f}(X_r^1, \ldots, X_r^n) = (f(X^1, \ldots, X^n))_r. \tag{3}$$

In fact, we have on using (1)

$$\begin{aligned}(f(X_r^1, X^2, \ldots, X^n))_r &= (\cup f(X_r^1, x^2, \ldots, x^n) \quad \text{for all} \quad x^i \in X^i, i \geqq 2)_r = \\ &= \vee (f(X_r^1, x^2, \ldots, x^n))_r = \\ &= \vee [\vee (f(x^1, x^2, \ldots, x^n) \quad \text{for} \quad x^1 \in X^1)_r \quad \text{for} \quad x^i \in X^i, i \geqq 2]_r = \\ &= \vee (f(x^1, x^2, \ldots, x^n) \quad \text{for} \quad x^i \in X^i)_r = (f(X^1, X^2, \ldots, X^n))_r\end{aligned}$$

whence the trivial conclusion is derived that, in $(f(X^1, \ldots, X^n))_r$, X^1 can be

---

[3]) We could write the definition of $\bar{f}$ as

$$\bar{f}(X_r^1, \ldots, X_r^n) = (f(X_r^1, \ldots, X_r^n))_r$$

if we mean by $f(Y^1, \ldots, Y^n)$ for subsets Y^i of A the set of all $f(y^1, \ldots, y^n)$ with $y^i \in Y^i$. This notation will be used below.

replaced by its closure X_r^1, and clearly the same holds for each $i=2, \ldots, n$. This establishes (3).

Next we verify the following two assertions concerning r-admissible operations.

Theorem 1. *If f is an r-admissible operation, then $\bar{f}$ is infinitely distributive over* $\vee$.

We have evidently

$$\bar{f}(\bigvee_\alpha X_r^\alpha, X_r^2, \ldots, X_r^n) = \bar{f}((\bigcup_\alpha X^\alpha)_r, X_r^2, \ldots, X_r^n) =$$

$$= (f(\bigcup_\alpha X^\alpha, X^2, \ldots, X^n))_r = (\bigcup_\alpha f(X^\alpha, X^2, \ldots, X^n))_r =$$

$$= \bigvee_\alpha f(X^\alpha, X^2, \ldots, X^n)_r = \bigvee_\alpha \bar{f}(X_r^\alpha, X_r^2, \ldots, X_r^n).$$

Theorem 2. *Let $(A; F)$ be an algebra and $(A_r; F, \wedge, \vee)$ the lattice-ordered algebra of its r-ideals. If*

$$\varphi(x^1, \ldots, x^k) = \psi(x^1, \ldots, x^k)$$

is an identity in A such that

1. *φ and ψ are built up from r-admissible operations $f \in F$;*

2. *each of φ and ψ contains every variable x^i at most once explicitly,*[4]

then

$$\bar{\varphi}(X_r^1, \ldots, X_r^k) = \bar{\psi}(X_r^1, \ldots, X_r^k)$$

is an identity in A_r.

For the proof observe that in view of (3) and since no repetition of the x^i may occur in any member of the identity, $\bar{\varphi}(X_r^1, \ldots, X_r^k)$ is just the r-ideal generated by all $\varphi(x^1, \ldots, x^k)$ with $x^i \in X^i$. On account of the given identity, this is nothing else than the r-ideal generated by all $\psi(x^1, \ldots, x^k)$ with $x^i \in X^i$ which is in turn equal to $\bar{\psi}(X_r^1, \ldots, X_r^k)$. This completes the proof.

It is not always necessary to consider in $(A_r; F, \wedge, \vee)$ all the operations in F, it is sometimes sufficient to take only a subfamily of F as the family of fundamental operations in A_r. For instance, in the case of submodules we usually disregard in A_r from multiplications by ring elements and from subtraction.

The following problem is open: Given an algebra $(B; G, \wedge, \vee)$, when is it obtained from an algebra $(A; F)$ with $G \subseteq F$ as an r-ideal system? Note that B is always a complete lattice such that every element of B is the join of compact elements.[5] It is also a hard problem to find conditions under which the isomorphy of r-ideal systems implies the isomorphy of the algebras from which they were constructed.

---

[4]) Note that the commutative or associative identity satisfies condition 2, but distributivity does not.

[5]) Recall that an element c of a lattice K is said to be compact if $c \leqq \bigvee_\alpha x_\alpha$ for some (infinite) set $\{x_\alpha\} \subseteq K$ implies that $c \leqq x_{\alpha_1} \vee \ldots \vee x_{\alpha_t}$ for a finite number of the x_α.

2. The additive theory

If we regard the r-ideals of $(A; F)$ as analogues of the ideals in ring theory, then we may try to develop an additive ideal theory on the pattern of the Lasker—Noether theory of ideals in commutative rings. The fact that we have a lattice-ordered algebra $(A_r; F, \wedge, \vee)$ at our disposal which has been obtained from an algebra $(A; F)$ is not relevant in our discussions. Therefore we shall make use of the lattice-theoretic point of view developed in different generalizations of commutative ideal theory.

In what follows let $(L; F)$ be a lattice-ordered algebra [which will play the role of $(A_r; F, \wedge, \vee)$] satisfying the following conditions:

(i) the binary operations $\wedge, \vee$ are in F and under them L is a complete lattice;

(ii) every $f \in F$ is isotone in each of its variables with the whole of L as monotony domain;

(iii) every $f \in F$ is infinitely distributive over $\vee$, i. e.

$$f(x_1, \dots, \bigvee_\alpha x_i^\alpha, \dots, x_n) = \bigvee_\alpha f(x_1, \dots, x_i^\alpha, \dots, x_n)$$

for all x_j in L.

Note that our hypotheses imply that L contains a greatest and a least element, and if g is an operation composed of operations $f \in F$, then it is likewise infinitely distributive over $\vee$.

In addition to (i)—(iii) we also assume:

(iv) there is an operator Φ which associates with each element $x \in L$ an element $\Phi(x)$ of L such that

$$x \leqq \Phi(x).$$

In most cases it is useful to assume that Φ is a closure operator, i. e.

$$x \leqq \Phi(y) \text{ implies } \Phi(x) \leqq \Phi(y) \qquad \text{for } x, y \in L,$$

or that it is linear in the sense that

$$\Phi(x \wedge y) = \Phi(x) \wedge \Phi(y) \qquad \text{for all } x, y \in L.$$

With the aid of operator Φ one is able to introduce special types of elements corresponding to prime elements and generalizations in ring theory.[6])

Let f be an n-ary operation on L, $n \geqq 2$. We are going to call an element p of L an *(f, Φ)-prime* if

(a) for each $i = 1, \dots, n$ and for all $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n \in L$ we have

$$f(x_1, \dots, x_{i-1}, p, x_{i+1}, \dots, x_n) \leqq p,$$

(b) if f' is an operation built up from the operation f only,[7] and if

$$f'(x_1, \dots, x_k) \leqq p,$$

then

$$x_j \leqq \Phi(p) \qquad \text{for some } j = 1, \dots, k.$$

[6]) Examples for Φ may be found in [1], Chapter XII, section 6.

[7]) If e. g. f denotes multiplication, then f' is multiplication with several factors.

In particular, if Φ is the identity operator and if f stands for multiplication, then we have the just the notion of prime elements.

An element $q \in L$ is called (f, Φ)*-primary* if it satisfies, in addition to (a) (with p replaced by q), also

(c) if f' is as in (b) and if

$$f'(x_1, \ldots, x_k) \leqq q,$$

then for each i, $1 \leqq i \leqq k$,

$$\text{either } x_i \leqq q \text{ or } x_j \leqq \Phi(q) \text{ for some } j \neq i.$$

Clearly, this notion corresponds to primary ideals in commutative rings if $\Phi(x)$ is the radical of x. Observe that according to our definition, *every* (f, Φ)*-primary element is necessarily* (f, Φ)*-prime.*

We shall consider intersections $a = x_1 \wedge \ldots \wedge x_k$ of elements $a \in L$ by means of $x_i \in L$ possessing certain properties. We call such an intersection *irredundant* if no x_i may be omitted so as to obtain the same intersection a. It is called *short* if it is irredundant and no subset of the x_i has an intersection which has again the same property as the x_i under consideration.

The following results generalize those in ring theory and those in lattice-ordered semigroups.[8])

Theorem 3. *Let Φ be a closure operator. Then the intersection*

$$p = p_1 \wedge \ldots \wedge p_k$$

of a finite number of (f, Φ)-prime elements p_i is (f, Φ)-prime again if and only if

$$\Phi(p) = \Phi(p_j) \quad \textit{for some } j.$$

If p is (f, Φ)-prime, then take f' built up from the operation f only such that in the argument of f' each of p_i occurs at least once. Then by (a) and the isotony of f we have

$$f'(p_1, \ldots, p_k) \leqq p_i \quad \text{for each } i.$$

Hence $f'(p_1, \ldots, p_k) \leqq p$ which implies by assumption $p_j \leqq \Phi(p)$ for some $j = 1, \ldots, k$. Now Φ being a closure operator, $\Phi(p_j) \leqq \Phi(p)$. But $p \leqq p_j$ implies the converse inequality, thus $\Phi(p_j) = \Phi(p)$, establishing necessity. For the sufficiency, let $\Phi(p) = \Phi(p_j)$ for some j. Condition (a) is satisfied by p, because

$$f(x_1, \ldots, p, \ldots, x_n) \leqq f(x_1, \ldots, p_l, \ldots, x_n) \leqq p_l$$

for every l. If f' is an operation as stated in (b), and if $f'(x_1, \ldots, x_k) \leqq p$, then $f'(x_1, \ldots, x_k) \leqq p_j$ too which implies $x_l \leqq \Phi(p_j) = \Phi(p)$ for some index l.

Theorem 4. *If Φ is a linear closure operator and if*

$$a = p_1 \wedge \ldots \wedge p_k = p_1' \wedge \ldots \wedge p_m'$$

are two short decompositions of $a \in L$ into intersections of (f, Φ)-primes, then $k = m$, and the elements $\Phi(p_1), \ldots, \Phi(p_k)$ are the same as $\Phi(p_1'), \ldots, \Phi(p_m')$, up to order.

[8]) Cf. [1] Chapter XII, sections 4 and 5.

We begin the proof again with an f' as described in (b) such that $f'(p_1, \dots, p_k)$ contains each p_j at least once. Then

$$f'(p_1, \dots, p_k) \leqq p_1 \wedge \dots \wedge p_k = a \leqq p'_j$$

implies $p_i \leqq \Phi(p'_j)$ for some $i = i(j)$. Therefore $\Phi(p_i) \leqq \Phi(p'_j)$. The same inference with p'_j replaced by p_i yields $\Phi(p'_l) \leqq \Phi(p_i)$ for some $l = l(i)$. Now $\Phi(p'_j \wedge p'_l) = \Phi(p'_j) \wedge \Phi(p'_l) = \Phi(p'_l)$ implies in view of the preceding theorem that $p'_j \wedge p'_l$ is again (f, Φ)-prime, consequently, $j = l$ and we have $\Phi(p_i) = \Phi(p'_j)$. The proof is completed.

Turning to (f, Φ)-primary elements, we have:

Theorem 5. *Let Φ be a closure operator. An irredundant intersection*

$$q = q_1 \wedge \dots \wedge q_k$$

of a finite number of (f, Φ)-primary elements q_i is likewise (f, Φ)-primary if and only if

$$\Phi(q) = \Phi(q_j) \quad \textit{for each } j.$$

Let q be (f, Φ)-primary. Then we have by (a)

$$f(q_1 \wedge \dots \wedge q_{j-1} \wedge q_{j+1} \wedge \dots \wedge q_k, q_j, \dots, q_j) \leqq q$$

for each j where the first argument of f is certainly not $\leqq q$ because of irredundancy. Thus $q_j \leqq \Phi(q)$ and hence $\Phi(q_j) \leqq \Phi(q)$. The converse inclusion also holds, hence $\Phi(q_j) = \Phi(q)$ for an arbitrary j. Conversely, let q satisfy $\Phi(q) = \Phi(q_j)$ for every j. Since (a) is obvious for q, assume $f'(x_1, \dots, x_m) \leqq q$ for some f' as in (b). If $x_i \not\leqq q$, then $x_i \not\leqq q_j$ for some j. Therefore $f'(x_1, \dots, x_m) \leqq q_j$ implies $x_l \leqq \Phi(q_j) = \Phi(q)$ for some $l \neq i$. Consequently, q is (f, Φ)-primary.

We shall need the following generalization of residuals. This is not the most general one which can be introduced here in the natural way, but this will suffice in our present case.

Given $a, b \in L$ and an n-ary operation f with $n \geqq 2$, let us consider the set of all $x \in L$ such that

$$f(x, b, \dots, b) \leqq a.$$

If this set is not void, then by completeness and infinite distributivity it contains a maximum element which we shall denote by $a :_f b$. That is to say, $a :_f b$ satisfies:

$$x \leqq a :_f b \quad \text{if and only if} \quad f(x, b, \dots, b) \leqq a.$$

It is readily seen that $\left(\bigwedge_\alpha a_\alpha\right) :_f b$ exists if and only if each of $a_\alpha :_f b$ exists and then

$$\left(\bigwedge_\alpha a_\alpha\right) :_f b = \bigwedge_\alpha (a_\alpha :_f b).$$

Theorem 6. *Let Φ be a linear closure operator and*

$$a = q_1 \wedge \dots \wedge q_k = q'_1 \wedge \dots \wedge q'_m$$

short decompositions of $a \in L$ into (f, Φ)-primary elements. Then $k = m$ and the elements $\Phi(q_1), \dots, \Phi(q_k)$ are, up to order, equal to the elements $\Phi(q'_1), \dots, \Phi(q'_m)$.

To begin with, observe that the elements $\Phi(q_1), \ldots, \Phi(q_k)$ are different, and so are the elements $\Phi(q_1'), \ldots, \Phi(q_m')$ by virtue of the assumption on Φ, Theorem 5 and shortness of decompositions. Pick out some maximal one amongst all $\Phi(q_i)$, $\Phi(q_j')$, say $\Phi(q_1)$, and form the residual

$$a:_f q_1 = q_1:_f q_1 \wedge \ldots \wedge q_k:_f q_1.$$

This exists on account of (a). Since $q_1 \not\equiv \Phi(q_i)$ for $i>1$ [otherwise $\Phi(q_1) \leqq \Phi(q_i)$], from

$$f(q_i:_f q_1, q_1, \ldots, q_1) \leqq q_i$$

and from the (f, Φ)-primary character of q_i we infer $q_i:_f q_1 \leqq q_i$. On the other hand, by (a), $q_i:_f q_1 \geqq q_i$ whence $q_i:_f q_1 = q_1$ for $i>1$. Therefore[9])

$$a:_f q_1 = q_2 \wedge \ldots \wedge q_k$$

which is $>a$ by irredundancy. Hence

$$a:_f q_1 = q_1':_f q_1 \wedge \ldots \wedge q_m':_f q_1$$

implies $q':_f q_1 > q_j'$ for some j. From

$$f(q_j':_f q_1, q_1, \ldots, q_1) \leqq q_j'$$

we get $q_1 \leqq \Phi(q_j')$, and so $\Phi(q_1) \leqq \Phi(q_j')$. This proves that the maximal ones among $\Phi(q_i)$ and those among $\Phi(q_j')$ coincide.

Next let $\Phi(q_1) = \Phi(q_1')$ be maximal. Since by Theorem 5 $q_1 \wedge q_1'$ is again (f, Φ)-primary with $\Phi(q_1 \wedge q_1') = \Phi(q_1)$, we can replace q_1 and q_1' by their intersection in the representations of a; or, otherwise expressed, $q_1 = q_1'$ may be assumed. Then

$$a:_f q_1 = q_2 \wedge \ldots \wedge q_k = q_2' \wedge \ldots \wedge q_m',$$

and an obvious induction completes the proof.

It is to be emphasized that the existence of decompositions of the elements of L into the intersection of (f, Φ)-prime or -primary elements is not true in general. For the case in which f is multiplication we refer to [1] Chapter XII where several examples for Φ are exhibited.

Reference

[1] L. FUCHS, *Partially ordered algebraic systems* (Oxford—London—New York—Paris, 1963).

(*Received December 23, 1964*)

[9]) From (a) it follows that $q_1:_f q_1$ is equal to the maximum element of L.

АВТОМАТЫ С ИЗОМОРФНЫМИ ПОЛУГРУППАМИ

Ф. ГЕЧЕГ и И. ПЕАК (Сегед)*)

К шестидесятилетию со дня рождения профессора Л. Кальмара

Предлагаемая работа посвящена изучению связей между циклическими автоматами, с одной стороны, и принадлежащими им некоторыми полугруппами, — с другой. Статья отчасти примыкает к работе [2], в которой рассматривались с этой точки зрения коммутативные циклические автоматы.

Помимо некоторых определений и обозначений, взятых из работы [2], мы пользуемся еще следующими определениями.

Автомат $A=(\mathfrak{A}, \mathfrak{X}, \delta)$ называется приведенным по входам, если для произвольных $x_1, x_2\ (\in \mathfrak{X})$ имеет место заключение

$$\forall_{a\in\mathfrak{A}} a[\delta(a, x_1)=\delta(a, x_2)]\Rightarrow x_1=x_2.$$

Как обычно, мы говорим, что автомат $A'=(\mathfrak{A}', \mathfrak{X}', \delta')$ является гомоморфным образом автомата $A=(\mathfrak{A}, \mathfrak{X}, \delta)$, если существует пара (φ, ψ) однозначных отображений $\varphi\colon \mathfrak{A}\to\mathfrak{A}'$ и $\psi\colon \mathfrak{X}\to\mathfrak{X}'$ множества $\mathfrak{A}$ на $\mathfrak{A}'$ и множества $\mathfrak{X}$ на $\mathfrak{X}'$, соответственно, для которых имеет место

$$\varphi(\delta(a, x))=\delta'(\varphi(a), \psi(x)) \qquad (a\in\mathfrak{A}, x\in\mathfrak{X}). \tag{1}$$

Если φ и ψ взаимно однозначны, то автоматы A и A' называются изоморфными и это обстоятельство обозначается через $A\approx A'$.

Мы условимся говорить, что автомат $A'=(\mathfrak{A}', \mathfrak{X}', \delta')$ является частичным $\mathfrak{X}$-гомоморфным образом некоторого автомата $A=(\mathfrak{A}, \mathfrak{X}, \delta)$, если найдутся некоторое непустое подмножество $\mathfrak{X}''\ (\subseteq\mathfrak{X})$ и однозначные отображения $\varphi\colon \mathfrak{A}\to\mathfrak{A}'$ и $\psi\colon \mathfrak{X}''\to\mathfrak{X}'$ множества $\mathfrak{A}$ на $\mathfrak{A}'$ и множества $\mathfrak{X}''$ на $\mathfrak{X}'$, соответственно, и для этих φ и ψ имеет место (1).

Мы говорим, что автомат $A=(\mathfrak{A}, \mathfrak{X}, \delta)$ $\mathfrak{X}$-изоморфно вложим в автомат $A'=(\mathfrak{A}', \mathfrak{X}', \delta')$, если найдутся взаимно однозначные отображения $\varphi\colon\mathfrak{A}\to\mathfrak{A}'$ и $\psi\colon\mathfrak{X}\to\mathfrak{X}'$ так, что φ отображает множество $\mathfrak{A}$ *на* $\mathfrak{A}'$, ψ отображает множество $\mathfrak{X}$ *в* $\mathfrak{X}'$ и для этих φ и ψ имеет место (1).

Известно, что для каждой полугруппы S существует автомат A, удовлетворяющий условию $S(A)\approx S$. Исходя из S, такой автомат можно сконструировать например следующим образом. Пусть $\bar{S}$ обозначает полугруппу S, если она имеет единицу, или получающуюся из S присоединением единич-

---

* F. Gécseg and I. Peák (Szeged)

ного элемента полугруппу, если S не содержит единицы. Возьмем некоторые множества $\mathfrak{A}$ и $\mathfrak{X}$, для которых имеет место $|\mathfrak{A}|=O(\bar{S})$ и $|\mathfrak{X}|=O(S)$ [1]). После этого мы сконструируем автомат $A_S=(\mathfrak{A}, \mathfrak{X}, \delta)$, функция переходов δ которого определяется формулой $\delta(a, x)=f(f^{-1}(a)g^{-1}(x))$, где f и g являются некоторыми взаимно однозначными отображениями полугруппы $\bar{S}$ на множество $\mathfrak{A}$ и полугруппы S на $\mathfrak{X}$, соответственно. Легко убедиться, что $S(A_S)\approx S$ и A_S является приведенным по входам автоматом.

Имеет место следующая

Теорема 1. *Каждый приведенный по входам циклический автомат A является частичным $\mathfrak{X}$-гомоморфным образом автомата $A_{S(A)}$.*

Действительно, пусть $A=(\mathfrak{A}, \mathfrak{X}, \delta)$ является автоматом, удовлетворяющим условиям теоремы и возьмем автомат $A_{S(A)}=(\mathfrak{A}', \mathfrak{X}', \delta')$, где $|\mathfrak{A}'| = =O(\overline{S(A)})$, $|\mathfrak{X}'|=O(S(A))$, $\delta'(a', x')=f(f^{-1}(a')g^{-1}(x'))$ и $f:\overline{S(A)}\to\mathfrak{A}'$ и $g: S(A)\to\mathfrak{X}'$ являются взаимно однозначными отображениями полугруппы $\overline{S(A)}$ на $\mathfrak{A}'$ и полугруппы $S(A)$ на $\mathfrak{X}'$, соответственно. Как уже отмечалось выше, полугруппы $S(A)$ и $S(A_{S(A)})$ будут изоморфными и поэтому в дальнейшем мы можем отождествить полугруппу $S(A_{S(A)})$ с полугруппой $S(A)$.

Через a_0 обозначается некоторый образующий элемент автомата A. Пусть отображение $\varphi: \mathfrak{A}'\to\mathfrak{A}$ сопоставляет каждому состоянию a' $(\in\mathfrak{A}')$ автомата $A_{S(A)}$ то состояние a $(\in\mathfrak{A})$ автомата A, для которого $a=a_0p$ $(p\in F(\mathfrak{X}))$ [2]) и при упомянутом отождествлении имеет место $C_\varrho[p]=f^{-1}(a')$. При этом пусть $\varphi(f(e))=a_0$, где через e обозначается единичный элемент полугруппы $\overline{S(A)}$. Легко видеть тогда, что φ является однозначным отображением множества $\mathfrak{A}'$ на $\mathfrak{A}$. Рассмотрим теперь множество $\mathfrak{X}''$ $(\subseteq\mathfrak{X}')$, состоящее из тех и только тех входных сигналов x' автомата $A_{S(A)}$, для которых найдется x $(\in\mathfrak{X})$, такое, что при упомянутом выше отождествлении имеет место $C_\varrho[x]=g^{-1}(x')$ и определим отображение $\psi: \mathfrak{X}''\to\mathfrak{X}$, сопоставляющее каждому входному сигналу x' $(\in\mathfrak{X}'')$ входный сигнал x $(\in\mathfrak{X})$ автомата A, удовлетворяющий условию $C_\varrho[x]=g^{-1}(x')$. Не трудно показать, что ψ является однозначным отображением множества $\mathfrak{X}''$ на $\mathfrak{X}$.

Осталось еще показать, что для этих отображений φ и ψ имеет место (1). Для этой цели рассмотрим произвольное состояние a' $(\in\mathfrak{A}')$ и произвольный входный сигнал x' $(\in\mathfrak{X}'')$ автомата $A_{S(A)}$. Тогда найдутся p $(\in F(\mathfrak{X}))$ и x $(\in\mathfrak{X})$ так, что $a'=f(C_\varrho[p])$, $x'=g(C_\varrho[x])$ и мы имеем

$$\varphi(\delta'(a', x'))=\varphi(f(f^{-1}(a')g^{-1}(x')))=\varphi(f(C_\varrho[p]C_\varrho[x]))=$$
$$=\varphi(f(C_\varrho[px]))=a_0px=\delta(a_0p, x)=\delta(\varphi(a'), \psi(x')),$$

что и требовалось доказать.

В частности, в коммутативном случае справедлива следующая

Теорема 2. *Каждый приведенный по входам коммутативный циклический автомат A $\mathfrak{X}$-изоморфно вложим в автомат $A_{S(A)}$.*

---

[1]) $|\mathfrak{A}|$ и $O(S)$ обозначает мощность множества $\mathfrak{A}$ и порядок полугруппы S, соответственно.

[2]) Везде в этой статье, через $F(\mathfrak{X})$ обозначается свободная полугруппа без единицы в алфавите $\mathfrak{X}$.

Действительно, достаточно показать, что при таких условиях отображения φ и ψ, определенные в доказательстве теоремы 1 взаимно однозначны. Предположим, что для некоторых $a_1', a_2' (\in \mathfrak{A}')$ имеет место $\varphi(a_1') = \varphi(a_2')$. Тогда $a_0 p_1 = a_0 p_2$ $(p_i \in F(\mathfrak{X}), f^{-1}(a_i') = C_\varrho[p_i]$ $(i = 1, 2))$ и так для произвольного $b (= a_0 r)$ $(r \in F(\mathfrak{X}))$ из-за коммутативности мы имеем $bp_1 = a_0 r p_1 = a_0 p_1 r = a_0 p_2 r = a_0 r p_2 = b p_2$, т. е. $p_1 \equiv p_2(\varrho(A))$. Отсюда уже видно, что φ является взаимно однозначным отображением. Легко получается, что и отображение ψ также будет взаимно однозначным.

Из теорем 1 и 2 непосредственно вытекает

Теорема 3. *Для произвольной полугруппы S пусть $\mathfrak{M}_S$ обозначает множество всех приведенных по входам циклических автоматов, принадлежащих полугруппе S. Если $\mathfrak{M}_S \neq \varnothing$, то существует автомат A_0, такой, что каждый автомат из $\mathfrak{M}_S$ является частичным $\mathfrak{X}$-гомоморфным образом автомата A_0. Далее, если S содержит единицу, то $A_0 \in \mathfrak{M}_S$ и если при этом полугруппа S коммутативна, тогда каждый автомат из $\mathfrak{M}_S$ $\mathfrak{X}$-изоморфно вложим в автомат A_0. В частности, в случае конечного S автомат A_0 (с точностью до изоморфизма) однозначно определен.*

Действительно, таким автоматом A_0 может служить автомат A_S.

Вопрос об обращении теоремы 1 решается отрицательно. Придавая, однако, этой теореме более полную формулировку, она уже оказывается обратимой. Об этом говорит следующее утверждение, выделяющее, по существу, из совокупности всех частичных $\mathfrak{X}$-гомоморфных образов A_S автоматы с полугруппой S.

Теорема 4. *Если для конечной полугруппы S некоторый приведенный по входам циклический автомат $A = (\mathfrak{A}, \mathfrak{X}, \delta)$ является частичным $\mathfrak{X}$-гомоморфным образом автомата A_S, при некоторой паре отображений (φ, ψ), то имеет место $S(A) \approx S$ тогда и только тогда, если (1) множество всех элементов $g^{-1}(x')$ $(x' \in \mathfrak{X}'')$ является системой образующих полугруппы S и (2) разбиение $\pi_{\varphi f}$ полугруппы $S \cap \bar{S}$, индуцируемое отображением $\varphi f \colon \bar{S} \to \mathfrak{A}$ не имеет уточнения, являющегося нетривиальным двусторонне стабильным разбиением*[3]) *полугруппы S.*[4])

К доказательству достаточности рассмотрим полугруппу S и предположим, что автомат $A = (\mathfrak{A}, \mathfrak{X}, \delta)$ удовлетворяет условиям теоремы. Покажем, что тогда $S(A) \approx S$. Пусть $A'' = A_S/(\varphi, \psi)$ обозначает фактор-автомат автомата A_S, состояниями и множеством входных сигналов которого являются все классы множества $\mathfrak{A}'$ по разбиению, индуцируемому отображением φ и множество $\mathfrak{X}''$, соответственно, а функция переходов δ'' автомата A'' действует обычным образом. Ясно, что имеет место $S(A'') \approx S(A)$. Следовательно, достаточно показать, что $S(A'') \approx S$.

[3]) Разбиение π полугруппы S называется правосторонне (двусторонне) стабильным, если отношение эквивалентности в S, индуцируемое разбиением π является правосторонне (двусторонне) стабильным в смысле Е. С. Ляпина [1]. Разбиение полугруппы S мы назовем тривиальным, если все классы эквивалентности в этом разбиении суть отдельные элементы.

[4]) Мы здесь пользуемся обозначениями теоремы 1.

Рассмотрим отображение $\chi: F(\mathfrak{X}'') \to S$, сопоставляющее каждому слову $p = x_1' x_2' \dots x_k'$ $(\in F(\mathfrak{X}''))$, элемент $s = g^{-1}(x_1')g^{-1}(x_2')\dots g^{-1}(x_k')$ $(\in S)$. Сразу видно, что χ будет гомоморфным отображением свободной полугруппы $F(\mathfrak{X}'')$ на полугруппу S. Далее, если взять еще $p' = x_1'' x_2'' \dots x_l''$ $(\in F(\mathfrak{X}''))$ и предполагать $\chi(p) = \chi(p')$, то для всех a' $(\in \mathfrak{A}')$ мы имеем $\delta''(a', p) = f(f^{-1}(a')g^{-1}(x_1')g^{-1}(x_2')\dots g^{-1}(x_k')) = f(f^{-1}(a')g^{-1}(x_1'')g^{-1}(x_2'')\dots g^{-1}(x_l'')) = \delta''(a', p')$, следовательно, $p \equiv p'(\varrho(A''))$. А это означает, что полугруппа $S(A'')$ изоморфна некоторой фактор-полугруппе фактор-полугруппы $F(\mathfrak{X})/(\chi)$ свободной полугруппы $F(\mathfrak{X}'')$ по отношению конгруентности, индуцируемому в $F(\mathfrak{X}'')$ отображением χ. Так как $F(\mathfrak{X}'')/(\chi) \approx S$, то полугруппа $S(A'')$ изоморфна некоторой фактор-полугруппе полугруппы S. Возьмем теперь слова q, q' $(\in F(\mathfrak{X}''))$, удовлетворяющие условию $q \equiv q'$ $(\varrho(A''))$ и пусть e обозначает единичный элемент полугруппы $\bar{S}$. Тогда из $\varphi(f(e\chi(q))) = \varphi(f(e\chi(q')))$ следует $\chi(q) \equiv \chi(q')$ $(\pi_{\varphi f})$, т. е. $\chi(q)$ и $\chi(q')$ содержатся в одном и том же классе полугруппы $S(=\bar{S} \cap S)$ по $\pi_{\varphi f}$. Итак, мы получим, что полугруппа $S(A'')$ изоморфна фактор-полугруппе полугруппы S, принадлежащее разбиение π которой является двусторонне стабильным уточнением разбиения $\pi_{\varphi f}$. С другой стороны, по условиям теоремы, разбиение $\pi_{\varphi f}$ полугруппы S имеет только тривиальное двусторонне стабильное уточнение, так, что мы получим $S(A'') \approx S$. Этим самым достаточность доказана.

К доказательству необходимости мы предположим, что из условий (1) и (2) не выполняется только (2). В этом случае разбиение $\pi_{\varphi f}$ полугруппы S имеет отличное от тривиального (единственное) максимальное двусторонне стабильное уточнение π. Мы будем показать, что полугруппа $S(A'')$ изоморфна фактор-полугруппе полугруппы S, индуцируемой разбиением π. Ясно, что для этой цели достаточно убедиться в том, что для всех q, q' $(\in F(\mathfrak{X}''))$ из $\chi(q) \equiv \chi(q')\,(\pi)$ всегда следует $q \equiv q'(\varrho(A''))$. Возьмем произвольные a_1', a_2' $(\in \mathfrak{A}')$, для которых имеет место $\varphi(a_1') = \varphi(a_2')$. Тогда $f^{-1}(a_1')$ и $f^{-1}(a_2')$ входят в один и тот же класс по $\pi_{\varphi f}$ в полугруппе S. Однако, разбиение π является двусторонне стабильным уточнением правосторонне стабильного разбиения $\pi_{\varphi f}$, поэтому из-за $\chi(q) \equiv \chi(q')\,(\pi)$ мы получаем $f^{-1}(a_1')\chi(q) \equiv f^{-1}(a_2')\chi(q')\,(\pi_{\varphi f})$, другими словами, $\varphi(\delta'(a_1', q)) = \varphi(\delta'(a_2', q'))$, т. е. $q \equiv q'$ $(\varrho(A''))$.

В дальнейшем, мы предположим, что не выполняется условие (1). Тогда полугруппа S', порожденная множеством всех элементов $g^{-1}(x')$ $(x' \in \mathfrak{X}'')$ будет истинной подполугруппой полугруппы S. Итак, аналогично предыдущему, можно показать, что либо полугруппа $S(A'')$ изоморфна некоторой фактор-полугруппе полугруппы S', либо имеет место $S(A'') \approx S'$, в зависимости от того, имеет ли разбиение полугруппы $S' \cap \bar{S}$, индуцируемое отображением φf отличное от тривиального двусторонне стабильное уточнение, или же нет. Этим завершается доказательство теоремы 4.

Пусть задан произвольный автомат A. Интересный вопрос, при каких условиях определяется (с точностью до изоморфизма) автомат A полугруппой $S(A)$. Задавать такие условия — то же самое, как выделять такие классы автоматов, в которых автоматы с изоморфными полугруппами и самы являются изоморфными. Следующее предложение дает некоторое частное решение этой проблемы.

Теорема 5. *Если для конечных множеств* $\mathfrak{A}$ *и* $\mathfrak{X}$ *имеет место* $|\mathfrak{A}|=|\mathfrak{X}|$, *то приведенные по входам коммутативные циклические автоматы* $A=(\mathfrak{A}, \mathfrak{X}, \delta)$ *и* $A'=(\mathfrak{A}, \mathfrak{X}, \delta')$ *будут изоморфными тогда и только тогда, если* $S(A)\approx S(A')$.

Необходимость очевидна. К доказательству достаточности прежде всего покажем, что $A\approx A_{S(A)}$. Действительно, из теоремы 1 работы [2] следует, что полугруппа $S(A)$ содержит единицу и так по теореме 2 мы имеем $|\mathfrak{A}|=O(S(A))$. Отсюда из-за $|\mathfrak{A}|=|\mathfrak{X}|$ получается $|\mathfrak{X}|=O(S(A))$. С другой стороны, мощность множества $\mathfrak{X}'$ входных сигналов автомата $A_{S(A)}$ совпадает порядком полугруппы $S(A)$ и так $|\mathfrak{X}|=|\mathfrak{X}'|$. Следовательно, отображение ψ, определенное в доказательстве теоремы 1 будет взаимно однозначным, и поэтому $A\approx A_{S(A)}$. Аналогичным образом получается $A'\approx A_{S(A)}$, откуда из $A_{S(A)}\approx A_{S(A')}$ вытекает $A\approx A'$.

Литература

[1] Е. С. Ляпин, *Полугруппы* (Москва, 1960).
[2] И. Пеак, Автоматы и полугруппы. II, *Acta Sci. Math.*, **26** (1965), 49–54.

(Поступило 26/I, 1965)

АВТОМАТЫ И ПОЛУГРУППЫ. II

И. ПЕАК (Сегед)*

Результаты настоящей заметки примыкают к исследованиям, приведенным в работах А. К. Флекка [1] и Р. Х. Ёмке [3].

Автоматом называется объект $A=(\mathfrak{A},\mathfrak{X},\delta)$, состоящий из двух непустых множеств $\mathfrak{A}, \mathfrak{X}$ и функции $\delta: \mathfrak{A}\times\mathfrak{X}\to\mathfrak{A}$. Для произвольного состояния $a(\in\mathfrak{A})$ и произвольного слова $p=x_1x_2\dots x_k$ $(\in F(\mathfrak{X}))$[1] мы полагаем $ap=a_k$ $(\in\mathfrak{A})$, если $\delta(a,x_1)=a_1$, $\delta(a_1,x_2)=a_2$, $\dots$, $\delta(a_{k-1},x_k)=a_k$. Автомат $A=(\mathfrak{A},\mathfrak{X},\delta)$ называется коммутативным, если для произвольных $a(\in\mathfrak{A})$, p, q $(\in F(\mathfrak{X}))$ имеет место $apq=aqp$. Состояние $a_0(\in\mathfrak{A})$ автомата $A=(\mathfrak{A},\mathfrak{X},\delta)$ мы назовем образующим элементом автомата A, если для каждого $a\,(\in\mathfrak{A})$ существует слово $p\,(\in F(\mathfrak{X}))$ так, что $a=a_0p$. Автомат называется циклическим, если он имеет образующий элемент, и — сильно связным, если каждое его состояние является образующим элементом этого автомата.

Пусть задан произвольный автомат $A=(\mathfrak{A},\mathfrak{X},\delta)$. Через $E(A)$ мы условимся обозначать полугруппу всех эндоморфизмов автомата A, т. е. множество всех однозначных отображений η: $\mathfrak{A}\to\mathfrak{A}$, удовлетворяющих условию $\eta(\delta(a,x))=\delta(\eta(a),x)$ $(a\in\mathfrak{A};\, x\in\mathfrak{X})$, если в этом множестве произведение понимается в обычном смысле, как произведение отображений. Если эндоморфизм $\eta(\in E(A))$ является взаимно однозначным отображением множества $\mathfrak{A}$ на себя, то η называется автоморфизмом автомата A. Группа всех автоморфизмов автомата A обозначается через $G(A)$. Следуя В. М. Глушкову [2], каждому автомату $A=(\mathfrak{A},\mathfrak{X},\delta)$ мы сопоставляем полугруппу $S(A)$, определенную как фактор-полугруппа $F(\mathfrak{X})/\varrho(A)$ свободной полугруппы $F(\mathfrak{X})$ без единицы в алфавите $\mathfrak{X}$ по отношению конгруентности

$$p\equiv q(\varrho(A))\Leftrightarrow \forall_{a\in\mathfrak{A}}\, a[ap=aq] \quad (p,q\in F(\mathfrak{X})),$$

индуцируемому автоматом A.[2])

Естественным образом возникает вопрос о том, какие связи можно установить между некоторым автоматом A, с одной стороны, и принадлежащими ему полугруппами $S(A)$, $E(A)$ и группой $G(A)$, — с другой.

В работе А. К. Флекка [1] рассматриваются коммутативные сильно связные автоматы и устанавливаются некоторые связи между такими авто-

---

* I. Peák (Szeged)

[1]) Через $F(\mathfrak{X})$ обозначается свободная полугруппа без единицы в алфавите $\mathfrak{X}$.

[2]) Относительно дальнейших определений мы ссылаемся на работы [2] и [4].

матами и их группами автоморфизмов.[3]) Настоящая заметка посвящена изучению коммутативных циклических автоматов. Нижеследующие результаты обобщают некоторые результаты работы А. К. Флекка [1], продолжают исследование, приведенное в работе Р. Х. Ёмке [3] и показывают, что *в алгебраической теории коммутативных циклических автоматов полугруппа эндоморфизмов играет аналогичную роль, как группа автоморфизмов — в теории коммутативных сильно связных автоматов.*

Теорема 1. *Если коммутативный автомат* $A=(\mathfrak{A}, \mathfrak{X}, \delta)$ *является циклическим, то полугруппы* $S(A)$ *и* $E(A)$ *будут изоморфными и имеет место* $O(E(A))=|\mathfrak{A}|$ [4]).

Доказательство. Возьмем некоторый коммутативный циклический автомат $A=(\mathfrak{A}, \mathfrak{X}, \delta)$. Прежде всего мы покажем, что $O(S(A))=|\mathfrak{A}|$. Пусть $a_0(\in\mathfrak{A})$ является образующим элементом автомата A. Рассмотрим произвольное состояние $a\,(\in\mathfrak{A})$ и пусть $a=a_0p\;(p\in F(\mathfrak{X}))$. Если при этом $a=a_0p'$ $(p'\in F(\mathfrak{X}))$, то для каждого $b\,(\in\mathfrak{A})$ мы получим $bp=a_0rp=a_0pr=a_0p'r=$ $=a_0rp'=bp'$, где $b=a_0r$ $(r\in F(\mathfrak{X}))$, т. е. $p\equiv p'(\varrho(A))$. Это означает, что можно определить однозначное отображение множества $\mathfrak{A}$ в полугруппу $S(A)$: $a\to C[p]$ $(a\in\mathfrak{A},\, p\in F(\mathfrak{X}),\, a=a_0p)$. Ясно, что это отображение отображает множество $\mathfrak{A}$ на полугруппу $S(A)$ и легко убедиться, что оно будет взаимно однозначным. Следовательно, мы получили, что $O(S(A))=|\mathfrak{A}|$.

Далее, так как автомат A является коммутативным, то для каждого слова $p\,(\in F(\mathfrak{X}))$ отображение $\eta_p\colon a\to ap$ $(a\in\mathfrak{A})$ будет эндоморфизмом автомата A и $\eta_p=\eta_q \Leftrightarrow p\equiv q(\varrho(A))\,(p, q\in F(\mathfrak{X}))$. Легко получается, что отображение $C[p]\to\eta_p$ является изоморфным отображением полугруппы $S(A)$ в полугруппу $E(A)$.

Осталось еще убедиться в том, что отображение $C[p]\to\eta_p$ отображает полугруппу $S(A)$ на полугруппу $E(A)$. Для этой цели мы покажем, что для каждого эндоморфизма $\eta(\in E(A))$ найдется слово $p(\in F(\mathfrak{X}))$, для которого имеет место $\eta=\eta_p$. Действительно, если $\eta(a_0)=a$ и $a=a_0p$ $(p\in F(\mathfrak{X}))$, то для произвольного состояния $b(=a_0r;\; r\in F(\mathfrak{X}))$ мы получаем $\eta(b)=$ $=\eta(a_0r)=\eta(a_0)r=ar=a_0pr=a_0rp=bp$, т. е. имеет место $\eta=\eta_p$. Этим теорема 1 полностью доказана.[5])

Из теоремы 1 и из следствия теоремы 4 Флекка [1] мы получим, что *каждый эндоморфизм произвольного коммутативного сильно связного автомата является автоморфизмом.*

Мы заметим, что из того, что для некоторого коммутативного автомата A имеет место $S(A)\approx E(A)$, не следует, что автомат A будет циклическим.

[3]) В упомянутой работе А. К. Флекк рассматривает автомат в некотором обобщенном смысле.

[4]) $O(S)$ обозначает порядок полугруппы S, а $|\mathfrak{A}|$ — мощность множества $\mathfrak{A}$.

[5]) В другой формулировке, как Б. Чакань обратил внимание автора, эта теорема по существу означает, что в одном из результатов Флекка [1], согласно которому произвольная коммутативная и транзитивная полугруппа отображений некоторого множества в себя совпадает со своим централизатором в полугруппе всех отображений данного множества в себя, предположение о транзитивности может быть ослаблено. При таком подходе теорема 2 получается в качестве частного случая теоремы 1.

Действительно, коммутативный автомат

	0	1	2	3	4
x	1	2	0	4	3

удовлетворяет этому условию, но не является циклическим.

С другой стороны, существует автомат $A=(\mathfrak{A}, \mathfrak{X}, \delta)$, порядок полугруппы $E(A)$ которого равен мощности множества $\mathfrak{A}$, но он не является циклическим. Таким автоматом будет например автомат

	0	1
x	1	1

Возникает также вопрос о том, существует ли коммутативный автомат A, удовлетворяющий одновременно условиям $S(A)\approx E(A)$ и $O(E(A))=|\mathfrak{A}|$, который не является циклическим.

Как следующий пример покажет, ответ на этот вопрос является положительным. Действительно, коммутативный нециклический автомат A, заданный таблицей переходов

	0	1	2	3
x	0	1	2	3
y	1	0	2	3
z	0	1	3	2

удовлетворяет одновременно данным двум условиям. Далее, легко убедиться и в том, что для этого автомата A имеет место $E(A)\approx G(A)$. Этим доказано также, что то утверждение Флекка [1], обобщением которого является теорема 1 и согласно которому, если коммутативный автомат $A=(\mathfrak{A}, \mathfrak{X}, \delta)$ сильно связен, то имеют место соотношения $G(A)\approx S(A)$ и $O(G(A))=|\mathfrak{A}|$, оказывается необратимым.

Теорема 2. *Пусть циклический автомат A имеет n состояний. Если полугруппа $E(A)$ коммутативна и $O(E(A))=n$, тогда автомат A также является коммутативным.*

Доказательство. Рассмотрим циклический автомат $A=(\mathfrak{A}, \mathfrak{X}, \delta)$ с образующим элементом $a_0(\in\mathfrak{A})$, возьмем произвольный эндоморфизм $\eta\in\in E(A)$ и определим отображение $\eta\to\eta(a_0)$ полугруппы $E(A)$ в $\mathfrak{A}$. Легко убедиться, что это отображение будет взаимно однозначным и из-за $O(E(A))=n$ оно отображает полугруппу $E(A)$ на множество $\mathfrak{A}$. Следовательно, для каждого $a\,(\in\mathfrak{A})$ найдется $\eta(\in E(A))$, такое, что $\eta(a_0)=a$.

Предположим теперь, что автомат A не является коммутативным. Тогда существуют $a\,(\in\mathfrak{A})$ и $p, q\,(\in F(\mathfrak{X}))$, для которых $apq\neq aqp$. Возьмем тот эндоморфизм $\eta\,(\in E(A))$, для которого $\eta(a_0)=a$. Тогда $\eta(a_0pq)=\eta(a_0)pq==apq$ и $\eta(a_0qp)=\eta(a_0)qp=aqp$. Отсюда из-за $apq\neq aqp$ мы получим $a_0pq\neq\neq a_0qp$. Возьмем теперь эндоморфизмы $\eta_1, \eta_2(\in E(A))$, удовлетворяющие условиям $\eta_1(a_0)=a_0p$ и $\eta_2(a_0)=a_0q$. Тогда $\eta_1\eta_2(a_0)=\eta_1(\eta_2(a_0))=\eta_1(a_0q)=$

$=\eta_1(a_0)q=a_0qp$ и $\eta_2\eta_1(a_0)=\eta_2(\eta_1(a_0))=\eta_2(a_0p)=\eta_2(a_0)p=a_0qp$. Следовательно, $\eta_1\eta_2(a_0)\neq\eta_2\eta_1(a_0)$, что противоречит предположениям теоремы. Этим теорема 2 доказана.

Из теоремы 1 и из теоремы 3 работы [4] сразу вытекает справедливость следующего предложения.

Теорема 3. *Если конечный коммутативный циклический автомат A является прямым произведением автоматов $A_1, \ldots, A_n$, то и полугруппа $S(A)$ будет прямым произведением полугрупп $S(A_1), \ldots, S(A_n)$.*

Если рассматривать прямое произведение специального вида, а именно, — $\mathfrak{X}$-прямое произведение[6]), тогда без ограничения конечности можно высказать следующее предложение.

Теорема 4. *Коммутативный циклический автомат A является $\mathfrak{X}$-прямым произведением автоматов тогда и только тогда, если полугруппа $S(A)$ является прямым произведением полугрупп.*

Доказательство этой теоремы получается при помощи модификации доказательства теоремы 6 из работы А. К. Флекка [1].

Действительно, рассмотрим коммутативный циклический автомат $A=(\mathfrak{A}, \mathfrak{X}, \delta)$ с образующим элементом $a_0(\in\mathfrak{A})$, полугуппа $S(A)$ которого является прямым произведением полугрупп. Тогда по теореме 1 мы имеем $S(A)\approx E(A)$ и $E(A)$ тоже является прямым произведением некоторых полугрупп S_1 и S_2, т. е. $E(A)\approx S_1\otimes S_2$. Без ограничения общности мы можем отождествить полугруппы $E(A)$ и $S_1\otimes S_2$. Возьмем теперь автомат $A_1=(S_1, \mathfrak{X}, \delta_1)$, функция переходов δ_1 которого определяется следующим образом: Для $s_1\ (\in S_1)$ и $x\ (\in\mathfrak{X})$ мы полагаем $\delta_1(s_1, x)=s_1x_1(\in S_1)$, где $x_1(\in S_1)$ является первой компонентой эндоморфизма $\eta_x(\in E(A);\ a\to\delta(a, x))$ в прямом разложении $E(A)\approx S_1\otimes S_2$. Таким же образом определяется автомат $A_2=(S_2, \mathfrak{X}, \delta_2)$.

Достаточно еще показать, что автомат $A'=(S_1\times S_2, \mathfrak{X}, \delta')$, являющийся $\mathfrak{X}$-прямым произведением автоматов A_1 и A_2, изоморфен исходному автомату A. Легко убедиться, что искомым изоморфизмом может служить отображение $\varphi\colon S_1\times S_2\to\mathfrak{A}$, сопоставляющее каждому состоянию (s_1, s_2) автомата A' состояние a автомата A, являющееся образом образующего элемента $a_0\ (\in\mathfrak{A})$ при эндоморфизме $(s_1, s_2)\ \ (\in S_1\otimes S_2)$. Действительно, по теореме 1 отображение φ является взаимно однозначным отображениием множества $S_1\times S_2$ на множество $\mathfrak{A}$. С другой стороны, мы имеем $\varphi(\delta'((s_1, s_2), x))=\varphi((\delta_1(s_1, x), \delta_2(s_2, x)))=\varphi((s_1x_1, s_2x_2))=\varphi((s_1, s_2)(x_1, x_2))=$
$=\varphi((s_1, s_2)\eta_x)=((s_1, s_2)\eta_x)(a_0)=\eta_x((s_1, s_2)(a_0))=\eta_x(\varphi(s_1, s_2))=\delta(\varphi((s_1, s_2)), x)$.

Пусть наоборот предполагается, что коммутативный циклический автомат $A=(\mathfrak{A}, \mathfrak{X}, \delta)$ является $\mathfrak{X}$-прямым произведением автоматов $A_1=(\mathfrak{A}_1, \mathfrak{X}, \delta_1)$ и $A_2=(\mathfrak{A}_2, \mathfrak{X}, \delta_2)$. Ясно,что тогда автоматы A_1 и A_2 также являются комму-

---

[6]) Мы говорим, что автомат $A=(\mathfrak{A}, \mathfrak{X}, \delta)$ является $\mathfrak{X}$-прямым произведением автоматов $A_i=(\mathfrak{A}_i, \mathfrak{X}, \delta_i)$ $(i=1, \ldots, n)$, если A изоморфен автомату $A'=(\mathfrak{A}', \mathfrak{X}, \delta')$, у которого $\mathfrak{A}'=\mathfrak{A}_1\times\ldots\times\mathfrak{A}_n$ и функция δ' определяется следующим образом: $\delta'((a_1, \ldots, a_n), x)=$ $=(\delta_1(a_1, x), \ldots, \delta_n(a_n, x))$ $(a_i\in\mathfrak{A}_i, x\in\mathfrak{X})$.

тативными и циклическими. Пусть a_0 $(\in\mathfrak{A})$ есть образующий элемент автомата A. Если отождествить автомат A с автоматом $A'=(\mathfrak{A}_1\times\mathfrak{A}_2, \mathfrak{X}, \delta')$, являющимся $\mathfrak{X}$-прямым произведением автоматов A_1 и A_2, то состояния $a_0^{(1)}(\in\mathfrak{A}_1)$ и $a_0^{(2)}(\in\mathfrak{A}_2)$, удовлетворяющие условию $a_0=(a_0^{(1)}, a_0^{(2)})$, будут образующими элементами автоматов A_1 и A_2, соответственно. Определим теперь следующее отображение: Каждому элементу $C[p]$ $(p\in F(\mathfrak{X}))$ полугруппы $S(A)$ мы сопоставляем тот элемент $(C_1[p_1], C_2[p_2])$ $(p_1, p_2\in F(\mathfrak{X}))$ прямого произведения $S(A_1)\otimes S(A_2)$, для которого имеет место $a_0^{(1)}p_1=a_1$, $a_0^{(2)}p_2=a_2$ и $a_0p=(a_1, a_2)$ $(a_i\in\mathfrak{A}_i)$. Не трудно убедиться, что это отображение является изоморфным отображением полугруппы $S(A)$ на полугруппу $S(A_1)\otimes S(A_2)$. Теорема доказана.

Исследуя взаимосвязь между автоматами и принадлежащими им группами автоморфизмов и полугруппами эндоморфизмов, естественным образом возникает вопрос о том, можно ли произвольную группу и произвольную полугруппу с единицей интерпретировать как группу автоморфизмов и полугруппу эндоморфизмов некоторого автомата, соответственно.

Теорема 4 Флекка [1] и теорема 1 настоящей заметки позволяют дать положительный ответ на этот вопрос в коммутативном случае.

Известно, что для каждой полугруппы S существует автомат A, полугруппа $S(A)$ которого изоморфна полугруппе S. Если например полугруппа S содержит единичный элемент, тогда такой автомат можно сконструировать следующим образом. Пусть X обозначает некоторую систему образующих элементов исходной полугруппы S. Возьмем произвольное множество $\mathfrak{A}$, мощность которого совпадает порядком полугруппы S и — множество $\mathfrak{X}$, мощность которого равна мощности множества X. Рассмотрим наконец взаимно однозначные отображения $f\colon \mathfrak{A}\to S$ множества $\mathfrak{A}$ на S и $g\colon \mathfrak{A}\to X$ множества $\mathfrak{X}$ на X, соответственно. Не трудно убедиться, что для автомата $A_S=(\mathfrak{A}, \mathfrak{X}, \delta)$, у которого функция переходов δ определяется формулой $\delta(a, x)=f^{-1}(f(a)g(x))$ $(a\in\mathfrak{A}; x\in\mathfrak{X})$, имеет место $S(A_S)\approx S$.

Используя эту конструкцию, доказывается следующая

Теорема 5. *Для каждой абелевой группы G существует автомат A, группа автоморфизмов $G(A)$ которого изоморфна заданной группе G. Далее, для каждой коммутативной полугруппы S с единицей найдется автомат A, полугруппа эндоморфизмов $E(A)$ которого изоморфна заданной полугруппе S.*

Доказательство. Действительно, рассмотрим произвольную абелеву группу G и (при некоторой системе образующих) сконструируем автомат A_G. Легко видеть, что автомат A_G будет коммутативным и сильно связным. Следовательно, из-за $S(A_G)\approx G$ из теоремы 4 Флекка [1] получается $G(A_G)\approx G$.

Далее, если задана произвольная коммутативная полугруппа S с единицей e, то автомат A_S (при произвольно выбранной системе образующих) будет коммутативным и циклическим. (Его образующим элементом может служить например прообраз единичного элемента e полугруппы S при отображении f.) Следовательно, из $S(A_S)\approx S$ и из теоремы 1 мы получим, $E(A_S)\approx S$, что и требовалось доказать.

Литература

[1] A. C. Fleck, Isomorphism groups of automata, *J. Assoc. Comp. Machinery*, **9** (1962), 469–476.
[2] В. М. Глушков, Абстрактная теория автоматов, *Успехи матем. наук*, **16:5 (101)** (1961), 3–62.
[3] R. H. Oehmke, On the structures of an automaton and its input semigroup, *J. Assoc. Comp. Machinery*, **10** (1963), 521–525.
[4] И. Пеак, Автоматы и полугруппы. I, *Acta Sci. Math.*, **25** (1964), 193–201.

(Поступило 26/X, 1964)

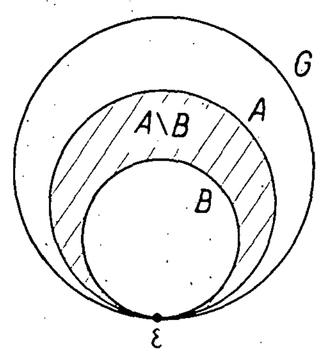
Ein Überdeckungssatz für endliche abelsche Gruppen im Zusammenhang mit dem Hauptsatz von Hajós

Von LADISLAUS RÉDEI in Szeged

Ladislaus Kalmár zum 60. Geburtstag gewidmet

Bekanntlich ist der Hauptsatz von HAJÓS[1]) für die endlichen abelschen Gruppen äquivalent mit einem von MINKOWSKI vermuteten berühmten Überdeckungssatz der endlichdimensionalen Euklidischen Räume, in dem es sich um eine Art Überdeckungen mit Würfeln handelt. Wir beweisen hier einen neuartigen Überdeckungssatz für die endlichen abelschen Gruppen, dessen wesentlichster Teil wieder mit dem Hauptsatz von HAJÓS äquivalent ist.

Es sei G eine endliche abelsche Gruppe mit dem Einselement ε. Eine (nichtleere) Differenzmenge $A \setminus B$ gebildet aus zwei Untergruppen $A \supset B$ von G nennen wir kurz eine *Sichel* (s. die Figur). Keine Sichel enthält ε, aber jedes von ε verschiedenes Element von G ist in mindestens einer Sichel, insbesondere nämlich gewiß in der maximalen Sichel $G \setminus \varepsilon$ enthalten.



Jede nichtleere Differenzmenge $A \setminus B$ von zwei Untergruppen A und B ist nach

$$A \setminus B = A \setminus (A \cap B)$$

gleich einer Sichel, wenn wir aber schlechtweg von einer Sichel $A \setminus B$ sprechen, so werde darin immer wie oben mit einbegriffen, daß A und B Untergruppen von G sind und $A \supset B$ ist.

Diese „Normalform" $A \setminus B$ einer Sichel ist eindeutig. Zum Beweis werde die Bemerkung vorausgeschickt, daß für Untergruppen U, V, W einer Gruppe mit $U \subseteqq V \cup W$ offenbar $U \subseteqq V$ oder $U \subseteqq W$ ist. Nun folgt aus der Relation

$$A \setminus B \subseteqq C \setminus D$$

zwischen zwei Sicheln zunächst (aus rein mengentheoretischen Gründen) $A \subseteqq B \cup C$, woraus (wegen $A \supset B$) nach obiger Bemerkung weiter $A \subseteqq C$ folgt. Im Fall der

[1]) G. HAJÓS, Über einfache und mehrfache Bedeckung des n-dimensionalen Raumes mit einem Würfelgitter, *Math. Zeitschr.*, **47** (1942), 427–467.

Gleichheit beider Sicheln folgt ähnlich $C \subseteqq A$, also muß $A=C$ und (wegen $A \supset B$ und $C \supset D$) auch $B=D$ sein. Damit ist die Behauptung bewiesen.

Den Index $(A:B)$ von B in A nennen wir auch den *Index* der Sichel $A \setminus B$. Nach dem soeben Bewiesenen ist also der Index einer Sichel eindeutig bestimmt.

Für Komplexe $K, K_1, \ldots, K_r$ von G mit $K = K_1 \cup \ldots \cup K_r$ sagen wir, daß $K_1, \ldots, K_r$ eine *Überdeckung* von K bilden. Die $K_1, \ldots, K_r$ nennen wir die Komponenten dieser Überdeckung. Zwei Überdeckungen, die voneinander nur in der Reihenfolge der Komponenten unterscheiden, werden oft als gleich angesehen. Wenn $K_1, \ldots, K_r$ paarweise disjunkt sind, so handelt es sich um die Klasseneinteilungen von K (in nichtleere Klassen $K_1, \ldots, K_r$).

Es läßt sich nach den Überdeckungen von $G \setminus \varepsilon$ mit Sicheln, d. h. nach Sicheln $A_1 \setminus B_1, \ldots, A_k \setminus B_k$ von der Eigenschaft

$$G \setminus \varepsilon = \bigcup_{i=1}^{k} (A_i \setminus B_i) \tag{1}$$

fragen. Das Produkt

$$\prod_{i=1}^{k} (A_i : B_i)$$

der Indizes der darin auftretenden Sicheln nennen wir den *Gesamtindex* dieser Überdeckung.

Wir verabreden, daß wir unter dem Durchschnitt $U_1 \cap \ldots \cap U_k$ von k Untergruppen einer gegebenen Gruppe im Fall $k=0$ diese Gruppe selbst verstehen. Dann gilt folgender leichter

Hilfssatz. *Sind $B_1, \ldots, B_k$ Untergruppen von G mit*

$$G \supset B_1 \supset B_1 \cap B_2 \supset \ldots \supset B_1 \cap \ldots \cap B_k = \varepsilon$$

und wird

$$A_i = (B_1 \cap \ldots \cap B_{i-1}) B_i \qquad (i=1, \ldots, k)$$

gesetzt, so existieren die Sicheln $A_i \setminus B_i$ $(i=1, \ldots, k)$ und bilden eine Überdeckung von $G \setminus \varepsilon$ vom Gesamtindex $(G:\varepsilon)$.

Zum Beweis setzen wir

$$D_i = B_1 \cap \ldots \cap B_i \qquad (i=0, \ldots, k).$$

Dann gelten

$$G = D_0 \supset D_1 \supset \ldots \supset D_k = \varepsilon$$

und

$$A_i = D_{i-1} B_i \qquad (i=1, \ldots, k).$$

Wegen $D_{i-1} \cap B_i = D_i$ ist

$$A_i \setminus B_i \supseteqq D_{i-1} \setminus D_i.$$

Da die rechte Seite nicht leer ist, muß $A_i \supset B_i$ sein, also ist $A_i \setminus B_i$ tatsächlich eine Sichel. Ferner folgt

$$\bigcup_{i=1}^{k} (A_i \setminus B_i) \supseteqq \bigcup_{i=1}^{k} (D_{i-1} \setminus D_i) = D_0 \setminus D_k = G \setminus \varepsilon,$$

weshalb die linke Seite (gleich $G \setminus \varepsilon$ ist, also) eine Überdeckung von $G \setminus \varepsilon$ liefert.

Schließlich ist die Faktorgruppe A_i/B_i nach dem ersten Isomorphiesatz (und wieder nach $D_{i-1}\cap B_i = D_i$) isomorph mit D_{i-1}/D_i, also ist $(A_i:B_i) = (D_{i-1}:D_i)$, woraus für den Gesamtindex dieser Überdeckung

$$\prod_{i=1}^{k}(A_i:B_i) = \prod_{i=1}^{k}(D_{i-1}:D_i) = (D_0:D_k) = (G:\varepsilon)$$

folgt. Damit ist der Hilfssatz bewiesen.

Wegen der Leichtigkeit dieses Beweises nennen wir eine Überdeckung von $G\setminus\varepsilon$ *trivial*, wenn sie von einer im Hilfssatz angegebenen Überdeckung sich höchstens in der Reihenfolge der Komponenten unterscheidet.

Nun wollen wir eine Untergruppe U von G mit zyklischer Faktorgruppe $G\setminus U$ *antizyklisch* nennen. Auch eine Sichel $A\setminus B$ nennen wir antizyklisch, wenn die Untergruppen A und B antizyklisch sind. (Es genügt, daß B antizyklisch ist, denn daraus folgt wegen $A\supset B$ das gleiche für A.) Es lautet der angekündigte

Überdeckungssatz. *Für eine endliche abelsche Gruppe G mit dem Einselement ε sind die Gesamtindizes der Überdeckungen von G mit antizyklischen Sicheln Vielfache der Ordnung $(G:\varepsilon)$ und diejenigen vom Gesamtindex $(G:\varepsilon)$ sind trivial.*

Zusatz. *Die zweite Hälfte dieses Satzes ist äquivalent mit der scharfen Form des Hauptsatzes von Hajós.*

Bemerkung. Bezüglich der scharfen Form des Hauptsatzes von Hajós s. man das Vorwort einer anderen Arbeit[2]). Die zweite Hälfte des Überdeckungssatzes können wir nur auf dem Umwege beweisen, daß wir den Zusatz beweisen. Durch einen direkten Beweis würde ein neuer Beweis des Hauptsatzes von Hajós entstehen, aber ein solcher Beweis ist uns nicht einmal im verhältnismäßig sehr einfachen Fall gelungen, wo die Invarianten von G gleiche Primzahlquadrate sind. (Unter den bekannten zahlreichen Äquivalenten des Haupsatzes von Hajós ist gegenwärtig nur dieser Hauptsatz selbst einem direkten Beweis zugänglich.)

Im Beweis bedienen wir uns der auch im § 1 der Arbeit[2]) eingeführten Bezeichnungen. (Eine einzige Abweichung von diesen Bezeichnungen wird sein, daß wir für den gruppentheoretischen Index $(A:B)$ weiter auch diese Bezeichnung behalten, wogegen wir hierfür in der Arbeit[2]) die Bezeichnung $O(A\setminus B)$ verwendet haben.)

Wir machen den Kunstgriff, daß wir den Satz für die Charaktergruppe $\mathfrak{G}$ von G statt G beweisen. Das ist durchaus gestattet, da G und $\mathfrak{G}$ miteinander isomorph sind. (Der Zweck dieses Übergehens von G auf $\mathfrak{G}$ wird uns bald klar.) Es sei eine Überdeckung

$$\mathfrak{G}\setminus\chi_1 = \bigcup_{i=1}^{k}(\mathfrak{A}_i\setminus\mathfrak{B}_i) \tag{2}$$

von $\mathfrak{G}\setminus\chi_1$ mit antizyklischen Sicheln $\mathfrak{A}_i\setminus\mathfrak{B}_i$ aus $\mathfrak{G}$ angegeben. Wir haben zu beweisen, daß für den Gesamtindex von (2) die Teilbarkeit

$$(\mathfrak{G}:\chi_1)\,\Big|\,\prod_{i=1}^{k}(\mathfrak{A}_i:\mathfrak{B}_i) \tag{3}$$

[2]) L. Rédei, Die neue Theorie der endlichen abelschen Gruppen und Verallgemeinerung des Hauptsatzes von Hajós, *Acta Math. Acad. Sci. Hung.* (im Erscheinen).

zutrifft und daß die Überdeckung (2) trivial ist, wenn in (3) „=" statt „|" gilt. Die zweite Behauptung — wir wiederholen es — werden wir so beweisen, daß wir ihre Äquivalenz mit der scharfen Form des Hauptsatzes von Hajós nachweisen. So werden Satz und Zusatz bewiesen.

Für eine Untergruppe X von G bezeichnen wir mit X' die Untergruppe von $\mathfrak{G}$ bestehend aus denjenigen Charakteren $\chi \in \mathfrak{G}$, für die stets $\chi(\xi)=1\,(\xi \in X)$ ist. Dual hierzu bezeichnet $\mathfrak{X}'$ für eine Untergruppe $\mathfrak{X}$ von $\mathfrak{G}$ die Untergruppe von G bestehend aus denjenigen Elementen $\xi \in G$, für die stets $\chi(\xi)=1$ $(\chi \in \mathfrak{X})$ ist. Man nenne die beiden zueinander inversen eineindeutigen Abbildungen $X \to X'$ und $\mathfrak{X} \to \mathfrak{X}'$ etwa die erste bzw. zweite Dirichletsche Abbildung. Die erste ist bekanntlich ein Antiisomorphismus des Untergruppenverbandes von G auf den von $\mathfrak{G}$, die zweite ist der hierzu duale Antiisomorphismus.

Es folgt, daß die in (2) auftretenden antizyklischen Untergruppen $\mathfrak{B}_1, \ldots, \mathfrak{B}_k$ von $\mathfrak{G}$ als

(4) $$\mathfrak{B}_i=\{\beta_i\}' \qquad (i=1, \ldots, k)$$

angesetzt werden können, wobei für $\beta_1, \ldots, \beta_k$ beliebige von ε verschiedene Elemente von G in Frage kommen. Da ferner $\mathfrak{A}_i$ jede echte Obergruppe von $\mathfrak{B}_i$ in $\mathfrak{G}$ bedeuten kann, läßt sich wegen (4)

(5) $$\mathfrak{A}_i = \{\beta_i^{e_i}\}' \qquad (i=1, \ldots, k)$$

setzen, wobei für $e_1, \ldots, e_k$ beliebige natürliche Zahlen mit

(6) $$e_i | o(\beta_i), \quad e_i > 1 \qquad (i=1, \ldots, k)$$

in Frage kommen.

Wir fassen die nach (6) existierenden Simplexe

(7) $$[\beta_i, e_i] \qquad (i=1, \ldots, k)$$

aus G ins Auge. Nach Hilfssatz 10 der Arbeit[2]) ist der Nullifikator $\mathfrak{N}([\beta_i, e_i])$ von $[\beta_i, e_i]$ gleich der Menge derjenigen Charaktere $\chi \in \mathfrak{G}$, die den Bedingungen $\chi(\beta_i)^{e_i}=1$ und $\chi(\beta_i) \neq 1$ genügen. Da die erste Bedingung als $\chi(\beta_i^{e_i})=1$ geschrieben werden kann, folgt daraus nach (4) und (5)

$$\mathfrak{N}([\beta_i, e_i]) = \mathfrak{A}_i \setminus \mathfrak{B}_i \qquad (i=1, \ldots, k),$$

also drückt sich die Annahme (2) in der Form

(8) $$\mathfrak{G} \setminus \chi_1 = \bigcup_{i=1}^{k} \mathfrak{N}([\beta_i, e_i])$$

aus.

Andererseits folgt aus (4), (5) und (6) das Bestehen von

(9) $$e_i = (\mathfrak{A}_i : \mathfrak{B}_i) \qquad (i=1, \ldots, k),$$

also lautet die Teilbarkeit (3) (wegen $(\mathfrak{G}:\chi_1)=(G=\varepsilon)$) als

(10) $$(G:\varepsilon) | e_1 \ldots e_k.$$

Zur ersten Hälfte des Satzes genügt es also zu beweisen, daß (10) eine Folgerung von (8) ist.

Zu diesem Zweck bezeichnen wir mit m eine natürliche Zahl von der Eigenschaft

$$(G:\varepsilon)|me_1\ldots e_k$$

und zeigen, daß im Gruppenring $\mathfrak{J}(G)$ die Gleichung

$$\frac{me_1\ldots e_k}{(G:\varepsilon)}\bar{G}=m\prod_{i=1}^{k}\overline{[\beta_i,e_i]} \tag{11}$$

gilt. Nach Hilfssatz 6 der Arbeit[2]) haben wir hierzu nur nachzuweisen, daß jeder Charakter $\chi\in\mathfrak{G}$ an beiden Seiten der Gleichung (11) den gleichen Wert annimmt. Für $\chi=\chi_1$ ist das wegen

$$\chi_1(\bar{G})=(G:\varepsilon),\quad \chi_1(\overline{[\beta_i,e_i]})=e_i\qquad (i=1,\ldots,k)$$

klar. Für $\chi\neq\chi_1$ (d. h. $\chi\in\mathfrak{G}\setminus\chi_1$) folgt die Behauptung aus $\chi(\bar{G})=0$ und (8).

Wegen des Auftretens des Zahlfaktors m auf der rechten Seite von (11) muß der Zahlfaktor auf der linken Seite durch m teilbar d. h. $e_1\ldots e_k$ durch den Nenner $(G:\varepsilon)$ teilbar sein. Damit ist die Teilbarkeit (10) d. h. die erste Hälfte des Überdeckungssatzes bewiesen.

Da (3) wegen (9) mit (10) übereinstimmt, setzen wir dementsprechend im weiteren Teil des Beweises auch

$$e_1\ldots e_k=(G:\varepsilon) \tag{12}$$

voraus. Um den Beweis auszuführen, haben wir vor allem zu erforschen, wie sich die Bedingungen in den Bestimmungsstücken β_i und e_i $(i=1,\ldots,k)$ ausdrücken, damit die Überdeckung (2) trivial ist. Das ist (s. den Hilfssatz) dann und nur dann der Fall, wenn nach passender Umnumerierung der Sicheln $\mathfrak{A}_i\setminus\mathfrak{B}_i$ $(i=1,\ldots,k)$

$$\mathfrak{G}\supset\mathfrak{B}_1\supset\mathfrak{B}_1\cap\mathfrak{B}_2\supset\ldots\supset\mathfrak{B}_1\cap\ldots\cap\mathfrak{B}_k=\varepsilon$$

und

$$\mathfrak{A}_i=(\mathfrak{B}_1\cap\ldots\cap\mathfrak{B}_{i-1})\mathfrak{B}_i\qquad (i=1,\ldots,k)$$

besteht. Mit Verwendung der zweiten Dirichletschen Abbildung drücken sich diese Bedingungen durch

$$\varepsilon\subset\mathfrak{B}_1'\subset\mathfrak{B}_1'\mathfrak{B}_2'\subset\ldots\subset\mathfrak{B}_1'\ldots\mathfrak{B}_k'=G$$

und

$$\mathfrak{A}_i'=\mathfrak{B}_1'\ldots\mathfrak{B}_{i-1}'\cap\mathfrak{B}_i\qquad (i=1,\ldots,k)$$

aus. Da aber nach (4) und (5) $\mathfrak{B}_i'=\{\beta_i\}$ und $\mathfrak{A}_i'=\{\beta_i^{e_i}\}$ gelten, verwandeln sich diese Bedingungen in

$$\varepsilon\subset\{\beta_1\}\subset\{\beta_1,\beta_2\}\subset\ldots\subset\{\beta_1,\ldots,\beta_k\}=G \tag{13}$$

und

$$\{\beta_i^{e_i}\}=\{\beta_1,\ldots,\beta_{i-1}\}\cap\{\beta_i\}\qquad (i=1,\ldots,k). \tag{14}$$

In (14) sind beide Seiten Untergruppen der zyklischen Gruppe $\{\beta_i\}$, also ist die Gleichung (14) gleichbedeutend mit der Gleichheit der Ordnung beider Seiten, d. h. (teils wegen (6_1)) mit

$$\frac{(\{\beta_i\}:\varepsilon)}{e_i}=\frac{(\{\beta_1,\ldots,\beta_{i-1}\}:\varepsilon)(\{\beta_i\}:\varepsilon)}{(\{\beta_1,\ldots,\beta_i\}:\varepsilon)}.$$

Für diese Gleichung darf $(\{\beta_1, \ldots, \beta_i\}:\varepsilon) = e_i(\{\beta_1, \ldots, \beta_{i-1}\}:\varepsilon)$ genommen werden, woraus folgt, daß das Gleichungssystem (14) mit dem System

$$(15) \qquad (\{\beta_1, \ldots, \beta_i\}:\varepsilon) = e_1 \ldots e_i \qquad (i=1, \ldots, k)$$

äquivalent ist. Da aber hieraus wegen (6_2) und (12) auch schon (13) folgt, haben wir gewonnen, daß die Überdeckung (2) dann und nur dann trivial ist, wenn nach passender Umnumerierung der Paare β_i, e_i die Gleichungen (15) erfüllt sind.

Nun ist aber (2) mit (12) zusammen nach dem Hauptlemma im § 1 der Arbeit[2]) gleichbedeutend damit, daß die schlichte Zerlegung

$$(16) \qquad G = [\beta_1, e_1] \circ \ldots \circ [\beta_k, e_k]$$

besteht. Also haben wir zur zweiten Hälfte unseres Satzes und zum Zusatz nur zu beweisen, daß die scharfe Form des Hauptsatzes von Hajós mit der Aussage äquivalent ist, daß aus (16) bei passender Umnumerierung der Faktoren die Bedingung (15) folgt.

Man bemerke noch, daß man dabei die obige Voraussetzung (6) fallen lassen darf, da diese in (16) schon mitenthalten ist. Aus der Existenz der Faktoren in (16) und aus der Definition der Simplexe folgt nämlich (6_2), ferner folgt aus (16) auch (6_1), da aus jeder Teilbarkeit

$$[\gamma, e] \mid G \qquad (\gamma \in G;\ o(\gamma) \geqq e > 1)$$

die weitere Teilbarkeit

$$e \mid o(\gamma)$$

folgt.

Um dieses zu beweisen, bemerke man, daß nach der Annahme ein Komplex K von G mit

$$G = [\gamma, e] \circ K$$

existiert. Dann ist

$$\bar{G} = \overline{[\gamma, e]}\bar{K}.$$

Hieraus folgt nach Multiplikation mit $\varepsilon - \gamma$

$$0 = (\varepsilon - \gamma^e)\bar{K}$$

d. h. $\gamma^e K = K$, also gilt sogar

$$\gamma^{(e, o(\gamma))} K = K.$$

Diese Gleichung ist mit der Existenz des obigen schlichten Produktes nur dann verträglich, wenn $(e, o(\gamma)) \geqq e$, d. h. $e \mid o(\gamma)$ ist. Damit ist die Behauptung bewiesen. Das bedeutet, daß es sich in (16) um alle möglichen Zerlegungen von G in ein schlichtes Produkt von Simplexen handelt.

Nunmehr besagt die scharfe Form des Hauptsatzes von Hajós (s. das Vorwort der Arbeit[2])), daß jede schlichte Zerlegung von G von der Form (16) klassisch ist, d. h. nach passender Numerierung der Faktoren alle schlichten Produkte

$$(17) \qquad P_i = [\beta_1, e_1] \circ \ldots \circ [\beta_i; e_i] \qquad (i=1, \ldots, k)$$

Gruppen sind. Wir zeigen, daß (15) dann und nur dann besteht, wenn alle schlichten Produkte (17) (existieren und) Gruppen sind, wodurch also die zweite Hälfte des Satzes und der Zusatz bewiesen sind.

Man sieht aus (17), daß stets, wenn $P_1, \ldots, P_k$ (existieren und) Gruppen sind, dann $P_i = \{\beta_1, \ldots, \beta_i\}$ und $(P_i : \varepsilon) = e_1 \ldots e_i$ $(i = 1, \ldots, k)$ gelten, woraus (15) folgt. Umgekehrt, wenn (15) besteht, so gilt

$$(\{\beta_1, \ldots, \beta_i\} : \{\beta_1, \ldots, \beta_{i-1}\}) = e_i$$

d. h.

$$\{\beta_1, \ldots, \beta_i\} = \{\beta_1, \ldots, \beta_{i-1}\} \circ [\beta_i, e_i],$$

woraus

$$\{\beta_1, \ldots, \beta_i\} = [\beta_1, e_1] \circ \ldots \circ [\beta_i, e_i] \qquad (i = 1, \ldots, k)$$

folgt. Da hiernach die Produkte (17) (existieren und) Gruppen sind, sind Satz und Zusatz bewiesen.

(Eingegangen am 18. Dezember 1964)

On certain representations of real numbers and on sequences of equivalent events

By ALFRED RÉNYI in Budapest

Dedicated to Professor L. Kalmár at the occasion of his 60th birthday

Introduction

In § 1 of this paper we shall deal with certain representations of real numbers. Let a_n $(n=0, 1, \dots)$ be an absolutely monotonic sequence of numbers, i. e. such that

(1) $$\Delta^k a_n > 0 \qquad (k=0, 1, \dots;\ n \geqq k),$$

where $\Delta^0 a_n = a_n$, $\Delta^1 a_n = \Delta a_n = a_{n-1} - a_n$ $(n \geqq 1)$, and

$$\Delta^k a_n = \Delta(\Delta^{k-1} a_n) \qquad (k \geqq 1, n \geqq k).$$

Let us also suppose that the sequence is *normed*, i. e.

(2) $$a_0 = 1,$$

and *regular*, i. e.

(3) $$a_n \to 0 \quad \text{and} \quad \Delta^n a_n \to 0 \qquad (n \to \infty).$$

Then every real $x \in (0, 1]$ admits a uniquely determined representation of the form

(4) $$x = \sum_{k=0}^{\infty} \Delta^k a_{n_k}$$

where the sequence of integers $1 \leqq n_0 < n_1 < n_2 < \dots$ depends on x.

This representation can also be written in the form

(5) $$x = \sum_{n=1}^{\infty} \varepsilon_n \Delta^{\varepsilon_1 + \dots + \varepsilon_{n-1}} a_n$$

where $\varepsilon_n = \varepsilon_n(x)$ equals 0 or 1; clearly $\varepsilon_n = 1$ if the number n occurs in the sequence n_k, and $\varepsilon_n = 0$ if not.

In § 2 we deal with the probability distribution of $n_k = n_k(x)$ provided that x is chosen at random with uniform distribution in (0, 1]. We shall show that the sequence of random variables n_k is then a Markov chain. In § 3 we deal with the joint probability distribution of the random variables ε_n $(n = 1, 2, \dots)$ defined above. We prove that if A_n denotes the random event that $\varepsilon_n(x) = 1$ then the events A_n

($n=1, 2, \ldots$) form a sequence of *equivalent* (symmetrically dependent) events, such that

(6) $$P(A_{m_1} A_{m_2} \ldots A_{m_k}) = \Delta^k a_k$$

for $1 \leqq m_1 < m_2 < \ldots < m_k$. (Here and in what follows $P(A)$ stands for the probability of the event A.)

In § 4 we show that the strong law of large numbers for equivalent events implies that for almost all x the limit

(7) $$\lim_{k\to\infty} \frac{k}{n_k(x)}$$

exists. (Previously in § 2 we obtain the weaker result that the distribution of k/n_k tends to a limit distribution.) On the other hand the above mentioned connection between equivalent events and the representation (4) or (5) leads to an effective construction of any sequence of equivalent events. A consequence of this is discussed in § 5. In § 6 we construct the corresponding measure preserving transformation to each sequence a_n, while in § 7 we discuss an example.

§ 1. Representation of real numbers by series of successive differences

We start with the following

Theorem 1. *Let a_n ($n=0, 1, \ldots$) be a normed, regular, absolutely monotonic sequence of real numbers. Then any real number $x \in (0, 1]$ can be represented in the form*

(1. 1) $$x = \sum_{k=0}^{\infty} \Delta^k a_{n_k}$$

where the increasing sequence of natural numbers n_k is uniquely determined by x.

Proof. Let n_0 be the first natural number such that

(1. 2) $$a_{n_0} < x;$$

such a number exists because $a_0 = 1$ and $a_n \to 0$. Let n_1 be the first natural number such that

(1. 3) $$a_{n_0} + \Delta a_{n_1} < x;$$

such a number n_1 exists because $\Delta a_n \to 0$. Moreover, by the definition of n_0 we have $a_{n_0} < x \leqq a_{n_0-1}$, i. e. $x - a_{n_0} \leqq \Delta a_{n_0}$, hence it follows $n_1 > n_0$. Similarly if $n_0, n_1, \ldots, n_r$ are already determined so that

(1. 4) $$\sum_{k=0}^{r} \Delta^k a_{n_k} < x \leqq \sum_{k=0}^{r-1} \Delta^k a_{n_k} + \Delta^r a_{n_r-1},$$

let n_{r+1} be the least natural number such that

(1. 5) $$\sum_{k=0}^{r+1} \Delta^k a_{n_k} < x.$$

It follows from (1.4) that

$$0<x-\sum_{k=0}^{r} \Delta^k a_{n_k} \leqq \Delta^{r+1} a_{n_r} \tag{1.6}$$

which implies — as by supposition $\Delta^{r+1}a_n$ is decreasing in n — that $n_{r+1}>n_r$.
Thus $n_{r+1} > r+1$. Therefore — using again the monotonicity of $\Delta^{r+1}a_n$ — it follows from (1.6) that

$$0<x-\sum_{k=0}^{r} \Delta^k a_{n_k} < \Delta^{r+1} a_{r+1}. \tag{1.7}$$

In view of the condition $\Delta^n a_n \to 0$ it follows that if the numbers n_k are determined by the algorithm described above, then (1.1) holds. This proves Theorem 1.

Let us note that according to a well-known theorem of F. HAUSDORFF [1] every normed absolutely monotonic sequence can be represented in the form

$$a_n = \int_0^1 t^n \, dF(t) \tag{1.8}$$

where $F(t)$ is non-decreasing on the closed interval [0, 1], is continuous from the left in the interior, and such that $F(0)=0$ and $F(1)=1$. Evidently,

$$\lim_{n\to\infty} a_n = F(1)-F(1-0),$$

thus condition $\lim a_n=0$ implies that $F(t)$ is continuous at $x=1$. We have further

$$\Delta^k a_n = \int_0^1 (1-t)^k t^{n-k} \, dF(t) \tag{1.9}$$

for $k=0, 1, \ldots$ and $n \geqq k$; thus in particular

$$\Delta^k a_k = \int_0^1 (1-t)^k \, dF(t). \tag{1.10}$$

Hence

$$\lim_{k\to\infty} \Delta^k a_k = F(+0).$$

Thus the condition of regularity $\lim \Delta^n a_n=0$ implies that $F(t)$ is continuous at $t=0$. Thus every normed, regular, absolutely monotonic sequence a_n can be represented in the form (1.8) where $F(t)$ is the distribution function of a probability distribution in the *open* interval (0, 1).

In view of formula (1.9) the representation (1.1) can be written in the form

$$x = \int_0^1 \left(\sum_{k=0}^{\infty} (1-t)^k t^{n_k-k} \right) dF(t). \tag{1.11}$$

Thus it follows that for every $x \in (0, 1]$ there is exactly one function $g(t)$ of the form

$$g(t) = \sum_{k=0}^{\infty} (1-t)^k t^{n_k - k}$$

with $1 \leqq n_0 < n_1 < n_2 < \ldots$, such that

$$x = \int_0^1 g(t)\, dF(t).$$

§ 2. Statistical theory of the difference-series representation of real numbers

Let x be a random variable, uniformly distributed in the interval (0, 1). Let us consider the representation of x in the form

$$x = \sum_{k=0}^{\infty} \Delta^k a_{n_k} \tag{2. 1}$$

where a_n is a given normed, regular, absolutely monotonic sequence. According to Theorem 1 the natural numbers $n_k = n_k(x)$ are uniquely determined by x; thus they are well defined random variables. We shall study now the probability laws governing the behaviour of these random variables. It is easy to see that if $n_1, \ldots, n_k$ are fixed, then x belongs to an interval of length $\Delta^{k+1} a_{n_k}$. It follows that denoting by $P(A|B)$ the conditional probability of the event A under condition B, we have

$$P(n_k = n \mid n_0 = m_0, n_1 = m_1, \ldots, n_{k-1} = m_{k-1}) = \frac{\Delta^{k+1} a_n}{\Delta^k a_{m_{k-1}}} \tag{2. 2a}$$

provided that $1 \leqq m_0 < m_1 < \ldots < m_{k-1} < n$. Thus the conditional distribution of n_k by given $n_0, \ldots, n_{k-1}$ depends on n_{k-1} only, that is the sequence of random variables n_k $(k = 0, 1, \ldots)$ is a Markov chain with the transition probabilities

$$P(n_k = n \mid n_{k-1} = m) = \frac{\Delta^{k+1} a_n}{\Delta^k a_m}. \tag{2. 2b}$$

As the probability on the right-hand side of (2. 2b) depends in general on k too, the Markov chain n_k is in general inhomogeneous. It is easy to see that the Markov chain is homogeneous if and only if $a_n = (1-p)^n$ $(n = 0, 1, \ldots)$ where $0 < p < 1$. In this particular case

$$P(n_k = n \mid n_{k-1} = m) = p(1-p)^{n-m-1}. \tag{2. 3}$$

This particular case corresponds to the representation of the real number x in the form

$$x = \sum_{k=0}^{\infty} p^k (1-p)^{n_k - k}. \tag{2. 4}$$

In this case if A_n denotes the event that n is contained in the sequence n_k then the events A_n $(n=1, 2, \dots)$ are independent and each has the probability $P(A_n)=p$. Especially if $p=\frac{1}{2}$ the representation (2. 4) reduces to

$$x = \sum_{k=0}^{\infty} \frac{1}{2^{n_k}} \quad \text{where} \quad 1 \leqq n_0 < n_1 < \dots, \tag{2.5}$$

by other words to the representation of x in the binary number system.

Let us return to the random variables n_k in the general case. The unconditional distribution of n_k can be determined as follows: As mentioned above if $n_0, n_1, \dots, n_k$ are fixed, then x belongs to an interval of length $\Delta^{k+1} a_{n_k}$. Now if only n_k is fixed, $n_k=n$, then the values of $n_0, n_1, \dots, n_{k-1}$ can be chosen in $\binom{n-1}{k}$ different ways; thus we have

$$P(n_k = n) = \binom{n-1}{k} \Delta^{k+1} a_n. \tag{2.6}$$

Especially in the case when $a_n = (1-p)^n$, we have

$$P(n_k = n) = \binom{n-1}{k} p^{k+1} (1-p)^{n-k-1} \tag{2.7}$$

i. e. $n_k - k - 1$ has a negative binomial distribution of order $k+1$.

In the general case it follows from (2. 6) and (1. 9) that

$$P(n_k = n) = \binom{n-1}{k} \int_0^1 (1-t)^{k+1} t^{n-k-1} \, dF(t) \tag{2.8}$$

for $n \geqq k+1$.

The distribution (2. 8) may be called a *mixed negative binomial distribution of order* $k+1$. The characteristic function of $\frac{n_k}{k+1}$ is

$$M\left(e^{\frac{iun_k}{k+1}}\right) = e^{iu} \int_0^1 \left(\frac{1-t}{1-te^{\frac{iu}{k+1}}} \right)^{k+1} dF(t). \tag{2.9}$$

(Here and in what follows M stands for "expectation".

We obtain by passing to the limit

$$\lim_{k\to\infty} M\left(e^{\frac{iun_k}{k+1}}\right) = \int_0^1 e^{\frac{iu}{1-t}} \, dF(t). \tag{2.10}$$

It follows that the probability distribution of $\frac{k+1}{n_k}$ tends to the distribution having the distribution function $1-F(1-z)$

$$\lim_{k\to\infty} P\left(\frac{k}{n_k} \leqq z\right) = 1 - F(1-z). \tag{2.11}$$

In the special case $a_n = (1-p)^n$ we have

$$F(t) = \begin{cases} 0 & \text{if} \quad t \leqq 1-p \\ 1 & \text{if} \quad t > 1-p, \end{cases}$$

thus (2.11) implies that in this case k/n_k tends in probability to p. This is of course well known, because

$$\frac{k}{n_k} = \frac{\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_{n_k}}{n_k} \tag{2.12}$$

and the (weak) law of large numbers applies to the independent random variables ε_n, each having the expectation p.

We shall show in the next paragraph that much more is true than (2.11): not only does the distribution of k/n_k tend for $k \to \infty$ to the limit distribution $1-F(1-z)$, but the random variables k/n_k themselves tend for $k \to \infty$ with probability 1 to a random variable $\varkappa$ having the distribution function $1-F(1-z)$.

Using the formula (2.8) we can of course compute all the moments of n_k. Especially we have

$$M(n_k) = (k+1) \int_0^1 \frac{dF(t)}{1-t}. \tag{2.13}$$

Thus the expectation of $\dfrac{n_k}{k+1}$ does not depend on k; it is finite if and only if the integral on the right of (2.13) is convergent, otherwise it is equal to $+\infty$.

§ 3. Connection with the theory of equivalent events

Let A_n denote the event that the natural number n is contained in the sequence $n_k(x)$ where x is a random variable, uniforly distributed in the interval (0, 1). We have evidently

$$P(A_n) = \sum_{k=0}^{n-1} P(n_k = n). \tag{3.1}$$

It follows from (2.8) that

$$P(A_n) = \int_0^1 \left(\sum_{k=0}^{n-1} \binom{n-1}{k} (1-t)^{k+1} t^{n-k-1} dF(t) = \int_0^1 (1-t)\, dF(t) \right. \tag{3.2}$$

for $n=1, 2, \ldots$. Before proceeding further we have to compute the r-step transition probabilities of the Markov-chain n_k. Clearly we have for $r \geqq 2$ and $n \geqq m+r$

$$P(n_{k+r} = n \mid n_k = m) = \frac{\Delta^{k+r+1} a_n}{\Delta^{k+1} a_m} \cdot \sum_{m < m_1 < \ldots < m_{r-1} < n} 1, \tag{3.3}$$

thus

$$P(n_{k+r} = n \mid n_k = m) = \binom{n-m-1}{r-1} \frac{\Delta^{k+r+1} a_n}{\Delta^{k+1} a_m}. \tag{3.4}$$

It follows that for $m<n$

$$P(n_{k+r}=n, n_k=m)=\binom{n-m-1}{r-1}\binom{m-1}{k}\Delta^{k+r+1}a_n. \tag{3.5}$$

Thus we have

$$P(A_m A_n)=\sum_{k=0}^{m-1}\sum_{l=k+1}^{k+n-m}\binom{n-m-1}{l-k-1}\binom{m-1}{k}\Delta^{l+1}a_n. \tag{3.6}$$

Taking (1.9) into account we obtain for $1\leqq m<n$

$$P(A_m A_n)=\int_0^1 (1-t)^2\,dF(t). \tag{3.7}$$

We shall show now that for any $r\geqq 1$ and for $1\leqq m_1<m_2<\ldots<m_r$ we have

$$P(A_{m_1}A_{m_2}\ldots A_{m_r})=\int_0^1 (1-t)^r\,dF(t). \tag{3.8}$$

The proof is essentially the same as for $r=2$. We obtain in the same way as (3.5) was shown — using that n_k is a Markov chain — that for $k_1<k_2<\ldots<k_r$, $m_1<m_2<\ldots<m_r$

$$P(n_{k_1}=m_1,\ldots,n_{k_r}=m_r)=\binom{m_1-1}{k_1}\prod_{j=1}^{r-1}\binom{m_{j+1}-m_j-1}{k_{j+1}-k_j-1}\cdot\Delta^{k_r+1}a_{m_r}. \tag{3.9}$$

Of course the probability (3.9) is positive only if $m_1\geqq k_1+1$ and $m_{j+1}-m_j\geqq k_{j+1}-k_j$ $(j=1,2,\ldots,r-1)$. From (3.9) one obtains (3.8) by means of the identity

$$P(A_{m_1}A_{m_2}\ldots A_{m_r})=\sum_{k_1<k_2<\ldots<k_r}P(n_{k_1}=m_1,\ldots,n_{k_r}=m_r). \tag{3.10}$$

As clearly

$$\int_0^1 (1-t)^r\,dF(t)=\Delta^r a_r \tag{3.11}$$

we have proved the following

Theorem 2. *Let A_n denote the event that the natural number n is contained in the sequence $\{n_k(x)\}$ defined by Theorem 1, where x is a random variable uniformly distributed in the interval $(0, 1)$. Then the events A_n $(n=1,2,\ldots)$ are equivalent, and one has for $1\leqq m_1<m_2<\ldots<m_r$ $(r=1,2,\ldots)$*

$$P(A_{m_1}A_{m_2}\ldots A_{m_r})=\Delta^r a_r. \tag{3.12}$$

Remark. Note that the sequence $W_r=\Delta^r a_r$ is absolutely monotonic too, because setting

$$G(t)=1-F(1-t+0) \tag{3.13}$$

we have

$$w_r=\int_0^1 t^r\,dG(t). \tag{3.14}$$

It is easy to see also that

$$\Delta^h w_r = \Delta^{r-h} a_r. \tag{3. 15}$$

Conversely let us be given a sequence of equivalent events B_n $(n=1, 2, \dots)$ in a probability space $[\Omega, \mathcal{A}, P]$ where Ω is a non empty set, the generic element of which will be denoted by ω, $\mathcal{A}$ is a σ-algebra of subsets of Ω and P a probability measure on $\mathcal{A}$. It is known (see [2], [3]) that there exists an $\mathcal{A}$-measurable function $\beta=\beta(\omega)$ on Ω — called the density of the sequence of events B_n — such that $0 \leqq \beta \leqq 1$ and for $r=1, 2, \dots$ and $m_1 < m_2 < \dots < m_r$ one has

$$P(B_{m_1} B_{m_2} \dots B_{m_r}) = \int_\Omega \beta^r \, dP. \tag{3. 16}$$

Let us consider first the case when $\beta=1$ on a set B of positive probability. Let β_n denote the indicator of the set B_n. It was shown in [3] that if $n_1 < n_2 < \dots < n_k$ $m_1 < m_2 < \dots < m$ and $n_i \neq m_j$, then

$$P(B_{n_1} B_{n_2} \dots B_{n_k} B_{m_1} B_{m_2} \dots B_{m_l}) = \int_\Omega \beta^k \beta_{m_1} \beta_{m_2} \dots \beta_{m_l} \, dP. \tag{3. 17}$$

It follows that

$$P\left(\prod_{n=r}^{\infty} B_n\right) = \int_B \prod_{r \leqq j < s} \beta_j \cdot dP. \tag{3. 18}$$

As (3. 18) holds for $s=r$ too (the empty product is equal to 1), we have

$$P\left(\prod_{n=r}^{\infty} B_n\right) = P(B).$$

Thus we obtain, putting $\prod_{n=1}^{\infty} B_n = B^*$,

$$P(B) = P(B^*) = P(BB^*).$$

This implies that the sets B and B^* are identical up to a set of P-measure 0. Let us denote now by $\bar{B}$ the complementary event of B, i. e. $\bar{B} = \Omega - B$. It follows that the events $A_n = \bar{B} B_n$ also are equivalent, and have the density α defined as follows:

$$\alpha(\omega) = \begin{cases} \beta(\omega) & \text{if} \quad \omega \in \bar{B}, \\ 0 & \text{if} \quad \omega \in B. \end{cases}$$

As a matter of fact we have

$$P(A_{m_1} A_{m_2} \dots A_{m_r}) = P(B_{m_1} B_{m_2} \dots B_{m_r}) - P(B) = \int_\Omega \alpha^k \, dP.$$

As $P(\alpha=1)=0$, we have shown that without restriction of generality one can suppose that $P(\beta=1)=0$.

Similarly one can suppose without restricting the generality that $P(\beta=0)=0$. As a matter of fact if C denotes the set on which $\beta=0$ and $0 < P(C) < 1$ then the set C is disjoint to all the sets B_n (up to a set of probability 0) and thus instead of the probability space $[\Omega, \mathcal{A}, P]$ we may consider the space $[\Omega, \mathcal{A}, P^*]$ where

$P^*(A)=\dfrac{P(A\bar{C})}{P(\bar{C})}$ and the events B_n will be equivalent with respect to this probability space too, with the same density β. Thus the case of an arbitrary sequence of equivalent events can be reduced to a sequence of equivalent events the density β of which is such that $P(\beta=0)=P(\beta=1)=0$. Let us call such a sequence a *regular* sequence of equivalent events. If A_n is a regular sequence of equivalent events with density β and if we put

$$w_k = P(A_{n_1}A_{n_2}\dots A_{n_k}) = \int_\Omega \beta^k\,dP$$

then clearly we have

$$\lim_{k\to\infty} w_k = 0 \quad \text{and} \quad \lim_{k\to\infty} \Delta^k w_k = 0.$$

Putting $a_k=\Delta^k w_k$, clearly $w_k=\Delta^k a_k$ and the sequence a_k is a normed regular absolutely monotonic sequence. Thus the events $\{A_n\}$ can be realized as the events connected with the representation of the random real number x uniformly distributed in (0, 1) in the form (1. 1), so that the event A_n is identified with the event that n is contained in the sequence n_k.

§ 4. The strong law of large numbers for the Markov chain n_k

We first give — to make this paper self-contained — a short proof of the following known result:[2])

Theorem 3. *Let A_n be an arbitrary sequence of equivalent events; let α_n denote the indicator of A_n and α the density of the sequence A_n. Then we have*

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\alpha_k=\alpha\right)=1. \tag{4. 1}$$

Proof. Let us consider the random variables

$$\delta_k = \alpha_k - \alpha \tag{4. 2}$$

and let us put for $k_1<k_2<\dots<k_r$, $r=1, 2, \dots$

$$P(A_{k_1}A_{k_2}\dots A_{k_r}) = w_r. \tag{4. 3}$$

It follows from (3. 17) that

$$M(\delta_{k_1}\delta_{k_2}\delta_{k_3}\delta_{k_4}) = \begin{cases} A & \text{if } k_1=k_2=k_3=k_4, \\ B & \text{if } k_1=k_2 \text{ and } k_3=k_4\neq k_1, \\ & \text{or if } k_1=k_3 \text{ and } k_2=k_4\neq k_1, \\ & \text{or if } k_1=k_4 \text{ and } k_2=k_3\neq k_1, \\ 0 & \text{otherwise,} \end{cases} \tag{4. 4}$$

---

[2]) Theorem 3 can also be deduced from BIRKHOFF's ergodic theorem.

where

(4. 5) $$A = w_1 - 4w_2 + 6w_3 - 3w_4$$

and

(4. 6) $$B = w_2 - 2w_3 + w_4.$$

This implies

(4. 7) $$M\left(\left(\frac{\delta_1 + \ldots + \delta_n}{n}\right)^4\right) = 0\left(\frac{1}{n^2}\right).$$

Thus the series

(4. 8) $$\sum_{n=1}^{\infty} \left(\frac{\delta_1 + \ldots + \delta_n}{n}\right)^4$$

is convergent with probability 1 and therefore (4. 1) holds.

In view of (2. 11) and the results of § 3 this implies that the following theorem holds:

Theorem 4. *If the sequence $n_k(x)$ is defined according to Theorem 1 then the limit*

(4. 9) $$\lim_{k\to\infty} \frac{k}{n_k(x)} = \varkappa(x)$$

exists for almost all x in $(0, 1)$; denoting by $\mu(A)$ the Lebesgue measure of the set A one has

(4. 10) $$\mu(\varkappa(x) \leqq y) = 1 - F(1-y) \quad \textit{for} \quad 0 \leqq y \leqq 1.$$

§ 5. Consequences for equivalent events

In the preceding § we applied the theory of equivalent sequences of events to prove the existence almost everywhere of the limit $\lim\limits_{k\to\infty} \dfrac{k}{n_k(x)}$. Conversely, our results lead to the proof of a property of equivalent events which seems not to be noticed up to now. This is expressed by

Theorem 5. *Let A_n $(n = 1, 2, \ldots)$ be a regular sequence of equivalent events. Let us set*

$$P(A_{n_1} A_{n_2} \ldots A_{n_k}) = w_k \qquad (n_1 < n_2 < \ldots < n_k;\ k = 1, 2, \ldots).$$

Denote by α_n the indicator of the event A_n and define the random variables ν_k as follows: ν_k is the least value of n such that $\alpha_1 + \alpha_2 + \ldots + \alpha_n = k$. By other words, ν_k denotes the index of the k-th event in the sequence of events A_n $(n = 1, 2, \ldots)$ which takes place. Then the random variables ν_k form a Markov chain with the transition probabilities

(5. 1) $$P(\nu_{k+1} = n \mid \nu_k = m) = \frac{\Delta^{n-k-1} w_n}{\Delta^{m-k} w_m}.$$

§ 6. The measure preserving transformation corresponding to a series of successive differences

To every representation (7) — i. e. to every normed, regular, absolutely monotonic sequence $\{a_n\}$ — there corresponds a measure preserving transformation T of the interval $(0, 1)$ defined as follows: If

$$x = \sum_{k=0}^{\infty} \Delta^k a_{n_k(x)} \tag{6. 1}$$

then

$$n_k(Tx) = n_{k+\varepsilon_1(x)}(x) = 1 \qquad (k=0, 1, \ldots). \tag{6. 2}$$

Clearly $1 \leqq n_0(Tx)$, because if $n_0(x)=1$ then $\varepsilon_1(x)=1$ and thus $n_0(Tx) = n_1(x)-1 \geqq 1$ and if $n_0(x) \geqq 2$ then $n_0(Tx) = n_0(x)-1 \geqq 1$; the inequality $n_{k+1}(Tx) > n_k(Tx)$ is evident. The inverse transformation $T^{-1}y$ can be defined as follows: $T^{-1}y$ is two-valued, namely if

$$y = \sum_{k=0}^{\infty} \Delta^k a_{n_k} \tag{6. 3}$$

then $T^{-1}y$ has the two values x_1 and x_2 where

$$x_1 = \sum_{k=0}^{\infty} \Delta^k a_{n_k+1}, \quad x_2 = a_1 + \sum_{k=1}^{\infty} \Delta^k a_{n_{k-1}+1}. \tag{6. 4}$$

Clearly if y belongs to the interval I_r defined by fixing the values of $n_0, n_1, \ldots, n_r$ in (6. 3) ($1 \leqq n_0 < n_1 < \ldots < n_r$) and having the length $\Delta^{r+1} a_{n_r}$ then x_1 belongs to an interval I_r' of length $\Delta^{r+1} a_{n_r+1}$ and x_2 to an interval I_r'' of length $\Delta^{r+2} a_{n_r+1}$. As

$$\Delta^{r+1} a_{n_r+1} + \Delta^{r+2} a_{n_r+1} = \Delta^{r+1} a_{n_r} \tag{6. 5}$$

it follows that denoting by $\mu(A)$ the Lebesgue measure of the set A one has

$$\mu(T^{-1} I_r) = \mu(I_r') + \mu(I_r'') = \mu(I_r). \tag{6. 6}$$

It follows from (6. 6) that Tx is measure preserving.

The transformation Tx can of course also be defined by

$$\varepsilon_k(Tx) = \varepsilon_{k+1}(x) \qquad (k=0, 1, \ldots). \tag{6. 7}$$

Thus T is equivalent to the shift transformation in the sequence-space $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n, \ldots)$.

It is easy to see that the transformation T is ergodic if and only if $a_n = q^n$ with $0 < q < 1$, because it follows from (6. 7) and Theorem 4 that

$$\varkappa(Tx) = \varkappa(x) \tag{6. 8}$$

and thus each level set of $\varkappa$ is an invariant set of T; thus T is ergodic if and only if $\varkappa$ is constant almost everywhere, i. e. if $a_n = q^n$. Especially in the case $a_n = 2^{-n}$, T is the well known transformation $Tx = (2x)$ where (Z) denotes the fractional part of Z.

§ 7. An example

As an example let us consider the sequence $a_n = \frac{1}{n+1}$ $(n=0, 1, 2, \dots)$. Evidently,

(7. 1)
$$\Delta^k a_n = \frac{1}{(n+1)\binom{n}{k}},$$

hence

(7. 2)
$$\Delta^n a_n = a_n = \frac{1}{n+1}.$$

Thus a_n is a normed, regular, absolutely monotonic sequence. Theorem 1 asserts for this case that every real number x with $0 < x \leqq 1$ has a unique representation of the form

(7. 3)
$$x = \sum_{k=0}^{\infty} \frac{1}{(n_k+1)\binom{n_k}{k}}$$

where the n_k are integers, $1 \leqq n_1 < n_2 < \dots$. The function $F(t)$ figuring in (1. 8) is in this example equal to t $(0 \leqq t \leqq 1)$. The transition probabilities (2. 2b) are in this example

(7. 4)
$$P(n_k = n \mid n_{k-1} = m) = \frac{(m+1)\binom{m}{k}}{(n+1)\binom{n}{k+1}}$$

and the distribution of n_k is given by

(7. 5)
$$P(n_k = n) = \frac{k+1}{n(n+1)} \quad \text{for} \quad n \geqq k+1.$$

Thus the random variables n_k have an infinite expectation. The equivalent events A_n can in this case be interpreted as the events of the following Pólya urn model: Let us consider an urn containing one white and one red ball. Let us draw one of the balls at random (each having the probability $\frac{1}{2}$ to be drawn) and put it back into the urn together with another ball of the same colour, then draw another ball from the urn which now contains 3 balls, each ball having the same probability to be drawn, put it back together with another ball of the same colour and continue this procedure indefinitely. Let A_n denote the event that at the n-th occasion a red ball has been drawn from the urn. Clearly in this interpretation a red ball is drawn the $k+1$-st time at the n_k-th drawing; the limit $\varkappa$ of k/n_k is in this case of course uniformly distributed in the interval $(0, 1)$.

References

[1] F. Hausdorff, Momentprobleme für ein endliches Intervall, *Math. Zeitschrift*, **16** (1923), 220—248.
[2] A. Rényi, On stable sequences of events, *Sankhya* (Ser. A), **25** (1963), 293—302.
[3] A. Rényi—P. Révész, A study of sequences of equivalent events as special stable sequences, *Publ. Math. Debrecen*, **10** (1963), 319—325.

(Received January 15. 1965)

Über ein Problem von S. B. Stetschkin

Von K. TANDORI in Szeged

Herrn Professor Ladislaus Kalmár zum 60. Geburtstag gewidmet

In dieser Note werden wir den folgenden Satz beweisen, der ein Problem von S. B. STETSCHKIN[1]) im positiven Sinne beantwortet.

Satz. *Es gibt eine trigonometrische Reihe*

$$\frac{a_0}{2}+\sum_{k=1}^{\infty}(a_k\cos kx+b_k\sin kx)$$

mit $a_k\to 0$, $b_k\to 0$ $(k\to\infty)$ *derart, daß für ihre Partialsummen* $s_n(x)$ *überall gilt:* $\underline{\lim}\, s_n(x)<\overline{\lim}\, s_n(x)$ $(n\to\infty)$.

Hilfssätze

Im folgenden bezeichnen wir mit $c_1, c_2, \ldots$ positive, absolute Konstanten.

Hilfssatz I. *Es seien* a *und* N *gegebene natürliche Zahlen. Dann gibt es ein trigonometrisches Polynom*

$$P(x)=P(a,N;x)=\sum_{k=\nu(a,N)}^{\mu(a,N)}(a_k\cos kx+b_k\sin kx)\qquad (N<\nu(a,N)<\mu(a,N))$$

mit den folgenden Eigenschaften: $|a_k|\leqq c_1$, $|b_k|\leqq c_1$,

$$|P(x)|\leqq c_2\max\left\{\frac{1}{a}\left(\frac{1}{x^2}+\frac{1}{(2\pi-x)^2}\right),a\right\}\qquad(-\infty<x<\infty),$$

und es gibt für jedes $x\in[-\pi/128a,\pi/128a]$ *Indizes* $p=p(x)$, $q=q(x)$ *derart, daß*

$$s_p(x)\geqq c_3a,\quad \text{und}\quad s_q(x)\leqq -c_3a\qquad (c_3\leqq 1),$$

wobei $s_n(x)$ *die* n*-te Partialsumme von* $P(x)$ *bezeichnet.*

---

[1]) Siehe П. Л. Улянов, Решенные и нерешенные проблемы теории тригонометрических и ортогональных рядов, *Успехи матем. наук*, **19 : 1 (115)** (1964), 3—69.

Wir setzen endlich

$$Q(x) = \sum_{i=0}^{68\varrho-1} \left(\frac{4c_8}{c_3}\right)^i Q_1\left(x-2i\frac{\pi}{128a}\right).$$

Auf Grund von (4) und (5) ist offensichtlich, daß $Q(x)$ allen Bedingungen des Hilfssatzes II genügt.

Beweis des Satzes

Es sei $(c_4 \leqq) a_1 < \ldots < a_i < \ldots$ eine Folge von natürlichen Zahlen, für die die Ungleichung

$$c_6(a_1 + \ldots + a_i) \leqq \frac{c_7}{2} a_{i+1} \qquad (i=1, 2, \ldots) \tag{6}$$

besteht und es sei $M_i = m(a_{i-1}^2, M_{i-1})$ $(i=1, 2, \ldots; M_0 = 0)$. Wir setzen

$$\sum_{i=1}^{\infty} \frac{1}{a_i} Q(a_i^2, m(a_{i-1}^2, M_{i-1}); x).$$

Auf Grund des Hilfssatzes II und (6) ist es klar, daß die Koeffizienten dieser trigonometrischen Reihe nach 0 streben und überall gilt:

$$\underline{\lim}\, s_n(x) = -\infty, \quad \overline{\lim}\, s_n(x) = \infty.$$

Damit haben wir unseren Satz bewiesen.

(Eingegangen am 11. März 1964)

Über ein Problem von S. B. Stetschkin

Von K. TANDORI in Szeged

Herrn Professor Ladislaus Kalmár zum 60. Geburtstag gewidmet

In dieser Note werden wir den folgenden Satz beweisen, der ein Problem von S. B. STETSCHKIN[1]) im positiven Sinne beantwortet.

Satz. *Es gibt eine trigonometrische Reihe*

$$\frac{a_0}{2}+\sum_{k=1}^{\infty}(a_k\cos kx+b_k\sin kx)$$

mit $a_k\to 0$, $b_k\to 0$ $(k\to\infty)$ *derart, daß für ihre Partialsummen* $s_n(x)$ *überall gilt:* $\underline{\lim}\, s_n(x)<\overline{\lim}\, s_n(x)$ $(n\to\infty)$.

Hilfssätze

Im folgenden bezeichnen wir mit $c_1, c_2, \ldots$ positive, absolute Konstanten.

Hilfssatz I. *Es seien a und N gegebene natürliche Zahlen. Dann gibt es ein trigonometrisches Polynom*

$$P(x)=P(a,N;x)=\sum_{k=\nu(a,N)}^{\mu(a,N)}(a_k\cos kx+b_k\sin kx)\qquad (N<\nu(a,N)<\mu(a,N))$$

mit den folgenden Eigenschaften: $|a_k|\leqq c_1$, $|b_k|\leqq c_1$,

$$|P(x)|\leqq c_2\max\left\{\frac{1}{a}\left(\frac{1}{x^2}+\frac{1}{(2\pi-x)^2}\right),a\right\}\qquad(-\infty<x<\infty),$$

und es gibt für jedes $x\in[-\pi/128a,\pi/128a]$ *Indizes* $p=p(x)$, $q=q(x)$ *derart, daß*

$$s_p(x)\geqq c_3a,\quad und\quad s_q(x)\leqq -c_3a\qquad (c_3\leqq 1),$$

wobei $s_n(x)$ *die n-te Partialsumme von* $P(x)$ *bezeichnet.*

---

[1]) Siehe П. Л. Ульянов, Решенные и нерешенные проблемы теории тригонометрических и ортогональных рядов, *Успехи матем. наук*, **19 : 1 (115)** (1964), 3—69.

Beweis. Es sei $N' = 96\cdot N\cdot a$. Da für $k=1, 2, 3, 4$

$$|l|\left(\frac{\pi}{N'}-\frac{\pi}{(N'+3ka)}\right) \leqq \frac{\pi}{8(N'+3ka)} \qquad \left(l=0, \pm 1, \ldots, \pm\frac{N'}{96a}\right)$$

gilt, so ergibt sich für $k=1, 2, 3, 4$

$$\sin(N'+3ka)x \geqq \sin\frac{\pi}{8}$$

(1)
$$\left(x\in\left[\frac{2l\pi}{N'}+\frac{\pi}{4N'}, \frac{(2l+1)\pi}{N'}-\frac{\pi}{4N'}\right]; -\frac{N'}{96a} \leqq 2l<2l+1 \leqq \frac{N'}{96a}\right),$$

$$\sin(N'+3ka)x \leqq -\sin\frac{\pi}{8}$$

$$\left(x\in\left[\frac{(2l-1)\pi}{N'}+\frac{\pi}{4N'}, \frac{2l\pi}{N'}-\frac{\pi}{4N'}\right]; -\frac{N'}{96a} \leqq 2l-1<2l \leqq \frac{N'}{96a}\right).$$

Für die a-te Fejérsche Kernfunktion

$$K_a(x) = \frac{1}{2}+\sum_{k=1}^{a}\left(1-\frac{k}{a+1}\right)\cos kx = \frac{1}{2(a+1)}\left(\frac{\sin(a+1)\frac{x}{2}}{\sin\frac{x}{2}}\right)^2$$

gelten die Abschätzungen

(2)
$$0 \leqq K_a(x) \leqq a, \quad K_a(x) \leqq \frac{\pi^2}{2a}\left(\frac{1}{x^2}+\frac{1}{(2\pi-x)^2}\right) \qquad (-\infty<x<\infty),$$

$$K_a(x) \geqq \frac{2}{\pi^2}a\left(-\frac{\pi}{2a} \leqq x \leqq \frac{\pi}{2a}\right).$$

Aus (1) und (2) folgt offensichtlich, daß das trigonometrische Polynom

$$P(x) = \sin(N'+3a)x\cdot K_a(x)-2\frac{\pi^2}{2\sin\frac{\pi}{8}}\sin(N'+6a)x\cdot K_a(x)+$$

$$+4\left(\frac{\pi^2}{2\sin\frac{\pi}{8}}\right)^2\sin(N'+9a)\left(x-\frac{\pi}{2N'}\right)\cdot K_a(x)-$$

$$-8\left(\frac{\pi^2}{2\sin\frac{\pi}{8}}\right)^2\sin(N'+12a)\left(x-\frac{\pi}{2N'}\right)\cdot K_a(x)$$

allen Bedingungen des Hilfssatzes I genügt.

Hilfssatz II. *Es seien $a(\geqq c_4)$ und M natürliche Zahlen. Dann gibt es ein trigonometrisches Polynom*

$$Q(x) = Q(a, M; x) = \sum_{k=n(a,M)}^{m(a,M)} (a_k \cos kx + b_k \sin kx) \qquad (M < n(a, M) < m(a, M))$$

mit den folgenden Eigenschaften: für jedes k sind $|a_k| \leqq c_5$, $|b_k| \leqq c_5$, es gilt die Abschätzung $|Q(x)| \leqq c_6 a$; und für jedes x gibt es Indizes $r = r(x)$, bzw. $s = s(x)$ derart, daß

$$S_r(x) \geqq c_7 a \quad \text{bzw.} \quad S_s(x) \leqq -c_7 a$$

gelten, wobei $S_n(x)$ die n-te Partialsumme von $Q(x)$ bezeichnet.

Beweis. Es sei ϱ eine natürliche Zahl mit

$$\frac{32 c_2^2}{\pi^2 \varrho^2} \leqq c_3, \tag{3}$$

wobei c_2 und c_3 die im Hilfssatz I erwähnten Konstanten bedeuten. Wir setzen

$$Q_1(x) = \sum_{1 \leqq i \leqq 2a/\varrho} P\left(a, N_i, x - i\varrho \frac{\pi}{a}\right),$$

wobei die Indizes N_i derart gewählt sind, daß $M < \nu(a, N_0) < \ldots < \mu(a, N_i) < < \nu(a, N_{i+1}) < \ldots$ gilt.

Ist $x \in \left[i_0 \varrho \frac{\pi}{a} - \frac{\pi}{128a}, i_0 \varrho \frac{\pi}{a} + \frac{\pi}{128a}\right]$ $(1 \leqq i_0 \leqq 2a/\varrho)$, dann gilt auf Grund des Hilfssatzes I und (3)

$$\sum_{i=1}^{i_0-1} \left| P\left(a, N_i; x - i\varrho \frac{\pi}{a}\right)\right| \leqq$$

$$\leqq \frac{c_2}{a} \sum_{i=1}^{i_0-1} \left(\frac{1}{\left(x - i\varrho \frac{\pi}{a}\right)^2} + \frac{1}{\left(2\pi - x + i\varrho \frac{\pi}{a}\right)^2} \right) \leqq \frac{16 c_2}{\pi^2 \varrho^2} a \leqq \frac{c_3}{2} a.$$

Daraus folgt, daß für $x \in \left[i\varrho \frac{\pi}{a} - \frac{\pi}{128a}, i\varrho \frac{\pi}{a} + \frac{\pi}{128a}\right]$ $(1 \leqq i \leqq 2a/\varrho)$ Indizes $\bar{r} = \bar{r}(x)$ und $\bar{s} = \bar{s}(x)$ existieren mit

$$\bar{S}_{\bar{r}}(x) \geqq \frac{c_3}{2} a, \quad \bar{S}_{\bar{s}}(x) \leqq -\frac{c_3}{2} a, \tag{4}$$

wobei $\bar{S}_n(x)$ die n-te Partialsumme von $Q_1(x)$ bezeichnet. Nach dem Hilfssatz I ergibt sich leicht mit $c_8 \geqq 1$

$$|Q_1(x)| \leqq \sum_{1 \leqq i \leqq 2a/\varrho} \left| P\left(a, N_i; x - i\varrho \frac{\pi}{a}\right)\right| \leqq 2c_2 \left(a + \frac{1}{a} \sum_{i=1}^{\infty} \frac{1}{\left(i\varrho \frac{\pi}{a}\right)^2}\right) \leqq c_8 a. \tag{5}$$

Wir setzen endlich

$$Q(x)=\sum_{i=0}^{68_{\varrho}-1}\left(\frac{4c_8}{c_3}\right)^i Q_1\left(x-2i\frac{\pi}{128a}\right).$$

Auf Grund von (4) und (5) ist offensichtlich, daß $Q(x)$ allen Bedingungen des Hilfssatzes II genügt.

Beweis des Satzes

Es sei $(c_4 \leqq) a_1 < \ldots < a_i < \ldots$ eine Folge von natürlichen Zahlen, für die die Ungleichung

$$c_6(a_1+\ldots+a_i) \leqq \frac{c_7}{2} a_{i+1} \qquad (i=1,2,\ldots) \tag{6}$$

besteht und es sei $M_i = m(a_{i-1}^2, M_{i-1})$ $(i=1,2,\ldots;\ M_0=0)$. Wir setzen

$$\sum_{i=1}^{\infty}\frac{1}{a_i} Q(a_i^2, m(a_{i-1}^2, M_{i-1}); x).$$

Auf Grund des Hilfssatzes II und (6) ist es klar, daß die Koeffizienten dieser trigonometrischen Reihe nach 0 streben und überall gilt:

$$\varliminf s_n(x) = -\infty, \quad \varlimsup s_n(x) = \infty.$$

Damit haben wir unseren Satz bewiesen.

(Eingegangen am 11. März 1964)

Sur les contractions de l'espace de Hilbert. X
Contractions similaires à des transformations unitaires

Par BÉLA SZ.-NAGY à Szeged et CIPRIAN FOIAŞ à Bucarest

Hommage à Monsieur L. Kalmár à son 60ième anniversaire

Le but de cette Note est de caractériser les contractions T de l'espace de Hilbert qui sont similaires à des transformations unitaires, et cela en termes de la fonction caractéristique de T. Grâce aux modèles fonctionnels des contractions basés sur la fonction caractéristique, on obtient de cette façon les modèles fonctionnels de toutes les contractions qui sont complètement non-unitaires, mais similaires à des transformations unitaires.

Rappelons que deux transformations, A et B, dans les espaces de Hilbert $\mathfrak{A}$ et $\mathfrak{B}$, bornées ou non, sont similaires lorsqu'il existe une transformation linéaire S de $\mathfrak{A}$ sur $\mathfrak{B}$, biunivoque et bicontinue, telle que $A = S^{-1}BS$.

A titre d'exemple, nous donnons une condition suffisante et nécessaire pour que la transformation

$$Af(x) = a(x)f(x) + i\int_0^x f(t)\,dt$$

dans l'espace $L^2(0, 1)$, où $a(x)$ est une fonction réelle et mesurable, soit similaire à une transformation autoadjointe.

1. Préliminaires

Soient T une contraction de l'espace de Hilbert $\mathfrak{H}$ et U la dilatation unitaire minimum de T, opérant dans un espace $\mathfrak{K} \supseteq \mathfrak{H}$. Soient

$$\mathfrak{K}_+ = \bigvee_0^\infty U^n\mathfrak{H}, \quad U_+ = U|\mathfrak{K}_+,$$

$$\mathfrak{L} = \overline{(U-T)\mathfrak{H}}, \quad \mathfrak{L}_* = \overline{(I-UT^*)\mathfrak{H}},$$

$$M(\mathfrak{L}) = \bigoplus_{-\infty}^{\infty} U^n\mathfrak{L}, \quad M_+(\mathfrak{L}) = \bigoplus_0^\infty U^n\mathfrak{L}, \quad M(\mathfrak{L}_*) = \bigoplus_{-\infty}^{\infty} U^n\mathfrak{L}_*, \quad M_+(\mathfrak{L}_*) = \bigoplus_0^\infty U^n\mathfrak{L}_*.$$

On a

(1. 1) $$\mathfrak{K}_+ = M_+(\mathfrak{L}_*) \oplus \mathfrak{K}_0 = M_+(\mathfrak{L}) \oplus \mathfrak{H},$$

$\mathfrak{K}_0$ étant un sous-espace de $\mathfrak{K}_+$, qui réduit U et par conséquent aussi U_+; soit

$$U_0 = U|\mathfrak{K}_0.$$

On a

$$T^* = U_+^*|\mathfrak{H}. \tag{1.2}$$

Les projections orthogonales de $\mathfrak{K}$ sur les sous-espaces

$$\mathfrak{H},\quad \mathfrak{K}_0,\quad M(\mathfrak{L}),\quad M(\mathfrak{L}_*)$$

soient désignées par

$$P_{\mathfrak{H}},\quad P_{\mathfrak{K}_0},\quad P^{\mathfrak{L}},\quad P^{\mathfrak{L}_*},$$

selon les cas. Les dernières trois projections permutent à U. Pour toutes ces notions et relations, *cf.* [VIII], n° 3 et [IX], n° 1.

On a la relation

$$P_{\mathfrak{K}_0}h = \lim_{n\to\infty} U^n T^{*n} h \qquad \text{pour tout } h\in\mathfrak{H}.\,[1] \tag{1.3}$$

Dans le cas où T n'a pas la valeur propre 0, on a aussi

$$\overline{P_{\mathfrak{K}_0}\mathfrak{H}} = \mathfrak{K}_0. \tag{1.4}$$

En effet, en cas contraire il existe un $k\in\mathfrak{K}_0$ tel que $k\perp\mathfrak{H}$ et $k\neq 0$. Puisque $k\in\mathfrak{K}_+\ominus\mathfrak{H} = M_+(\mathfrak{L})$, k admet un développement orthogonal

$$k = \sum_0^\infty U^n l_n \qquad \left(l_n\in\mathfrak{L},\ \sum_0^\infty \|l_n\|^2 = \|k\|^2\right).$$

Comme $k\neq 0$, il y a des coefficients l_n qui sont différents de 0: soit l_ν le premier d'entre eux. On a alors

$$U_+^{*\nu+1}k = U_+^* l_\nu + l_{\nu+1} + U_+ l_{\nu+2} + U_+^2 l_{\nu+3} + \dots. \tag{1.5}$$

Puisque $U_+^* l_\nu \in U_+^*\mathfrak{L} = U_+^*\,\overline{(U_+ - T)\mathfrak{H}} \subseteq \overline{(I-T^*T)\mathfrak{H}} \subseteq \mathfrak{H}$ (voir (1.2)) le développement (1.5) correspond à la décomposition orthogonale

$$\mathfrak{K}_+ = \mathfrak{H}\oplus\mathfrak{L}\oplus U\mathfrak{L}\oplus U^2\mathfrak{L}\oplus\dots,$$

d'où

$$\|U_+^{*\nu+1}k\|^2 = \|U_+^* l_\nu\|^2 + \|l_{\nu+1}\|^2 + \|l_{\nu+2}\|^2 + \|l_{\nu+3}\|^2 + \dots.$$

D'autre part, puisque $k\in\mathfrak{K}_0$, on a $U_+^{*\nu+1}k = U_0^{-\nu-1}k$, d'où

$$\|U_+^{*\nu+1}k\|^2 = \|k\|^2 = \|l_\nu\|^2 + \|l_{\nu+1}\|^2 + \|l_{\nu+2}\|^2 + \dots.$$

---

[1]) *Cf.* [VII], n° 1. La convergence $U^nT^{*n}h \to h_0$ s'ensuit de la relation

$$\|U^nT^{*n}h - U^mT^{*m}h\|^2 = \|T^{*m}h\|^2 - \|T^{*n}h\|^2 \qquad (h\in\mathfrak{H};\ m<n)$$

et de ce que les valeurs $\|T^{*n}h\|^2$ forment une suite non-croissante, donc convergente. On a

$$h - h_0 = \lim_n (I - U^nT^{*n})h = \lim_n \sum_{k=0}^{n-1} U^k(I-UT^*)T^{*k}h \in M_+(\mathfrak{L}_*).$$

D'autre part, $h_0 \perp M_+(\mathfrak{L}_*)$ parce que, pour tout ν fixé, $U^nT^{*n}h \perp U^\nu(I-UT^*)T^{*\nu}h'$ $(h, h'\in\mathfrak{H})$ dès que $n\geqq\nu+1$, conséquence de ce que $\mathfrak{H}\perp U^\mu\mathfrak{L}_*$ pour $\mu = -1, -2, \dots$. Donc $h_0 = P_{\mathfrak{K}_0}h$.

Il en résulte $\|U_+^* l_v\|^2 = \|l_v\|^2$, donc

(1. 6) $$((I_{\mathfrak{K}_+} - U_+ U_+^*) l_v, l_v) = 0.$$

Or, $I_{\mathfrak{K}_+} - U_+ U_+^*$ étant égal à la projection de $\mathfrak{K}_+$ sur $\mathfrak{K}_+ \ominus U_+ \mathfrak{K}_+ = \mathfrak{L}_*$ (*cf.* la première des relations (1.1)), (1. 6) veut dire que $l_v \perp \mathfrak{L}_*$. En vertu de la relation

$$\mathfrak{H} \oplus \mathfrak{L} = U\mathfrak{H} \oplus \mathfrak{L}_* \quad (\textit{cf.}\ [\mathrm{V}],\ \text{formule}\ (17))$$

cela entraîne que

$$l_v = Uh \quad (h \in \mathfrak{H}, h \neq 0), \quad P_{\mathfrak{H}} l_v = P_{\mathfrak{H}} Uh = Th.$$

Puisque $\mathfrak{L} \perp \mathfrak{H}$, on à $P_{\mathfrak{H}} l_v = 0$, donc $Th = 0$. Cela contredit notre hypothèse que T n'a pas la valeur propre 0. Donc (1. 4) subsiste.

Nous aurons besoin aussi du suivant simple

Lemme. *Soient*

$$\mathfrak{R} = \mathfrak{A} \oplus \mathfrak{B}, \quad \mathfrak{R} = \mathfrak{X} \oplus \mathfrak{Y}$$

deux décompositions d'un espace de Hilbert $\mathfrak{R}$. *Désignons par* $P_{\mathfrak{A}}$ *et* $P_{\mathfrak{B}}$ *les projections orthogonales dans* $\mathfrak{R}$ *sur* $\mathfrak{A}$ *et* $\mathfrak{B}$. *Supposons que* $P_{\mathfrak{A}}$ *applique* $\mathfrak{X}$ *sur* $\mathfrak{A}$ *de manière biunivoque et bicontinue.* $P_{\mathfrak{B}}$ *applique alors* $\mathfrak{Y}$ *sur* $\mathfrak{B}$ *de la même manière. Plus précisément: lorsqu'on a*

(1. 7) $$P_{\mathfrak{A}} \mathfrak{X} = \mathfrak{A} \quad \textit{et} \quad \|P_{\mathfrak{A}} x\| \geqq c\|x\| \quad \textit{pour tout} \quad x \in \mathfrak{X}$$

avec une constante positive c, on a aussi

(1. 8) $$P_{\mathfrak{B}} \mathfrak{Y} = \mathfrak{B} \quad \textit{et} \quad \|P_{\mathfrak{B}} y\| \geqq c\|y\| \quad \textit{pour tout} \quad y \in \mathfrak{Y}.$$

Démonstration. Puisque $\|P_{\mathfrak{A}} x\|^2 \geqq c^2\|x\|^2 = c^2[\|P_{\mathfrak{A}} x\|^2 + \|P_{\mathfrak{B}} x\|^2]$, on a $C^2\|P_{\mathfrak{A}} x\|^2 \geqq \|P_{\mathfrak{B}} x\|^2$ avec $C = \sqrt{1-c^2}/c$. Donc les hypothèses (1. 7) assurent que la formule

$$A(P_{\mathfrak{A}} x) = P_{\mathfrak{B}} x \qquad (x \in \mathfrak{X})$$

définit une transformation linéaire A de $\mathfrak{A}$ dans $\mathfrak{B}$, bornée par C. A^* sera alors une transformation linéaire de $\mathfrak{B}$ dans $\mathfrak{A}$, bornée aussi par C. Puisque le graphique $\{a \oplus Aa : a \in \mathfrak{A}\}$ de A dans $\mathfrak{R} = \mathfrak{A} \oplus \mathfrak{B}$ est égale à $\{P_{\mathfrak{A}} x \oplus P_{\mathfrak{B}} x : x \in \mathfrak{X}\}$, donc à $\mathfrak{X}$, son complément orthogonal $\{-A^* b \oplus b : b \in \mathfrak{B}\}$ sera égal à $\mathfrak{Y}$. Cela veut dire que $P_{\mathfrak{B}} \mathfrak{Y} = \mathfrak{B}$ et que $P_{\mathfrak{A}} y = -A^* P_{\mathfrak{B}} y$ pour tout $y \in \mathfrak{Y}$. Il s'ensuit que

$$\|P_{\mathfrak{A}} y\| \leqq C\|P_{\mathfrak{B}} y\|, \quad \|y\|^2 = \|P_{\mathfrak{A}} y\|^2 + \|P_{\mathfrak{B}} y\|^2 \leqq (1+C^2)\|P_{\mathfrak{B}} y\|^2 = \frac{1}{c^2}\|P_{\mathfrak{B}} y\|^2,$$

ce qui achève la démonstration des relations (1. 8).

2. Critère général de similitude à une transformation unitaire

1. Soit la contraction T de $\mathfrak{H}$ similaire à une transformation unitaire V, donc $T = S^{-1} V S$. Dans ce cas T^{-1} et T^{*-1} existent au sens strict et on a

$$T^{*-n} = S^* V^n S^{*-1}, \quad \|T^{*-n}\| \leqq \|S^*\|\,\|S^{*-1}\| = \frac{1}{c}, \quad \text{donc} \quad \|T^{*n} h\| \geqq c\|h\|$$

pour tout $h \in \mathfrak{H}$.

En vertu de (1. 3) on a

$$\|P_{\mathfrak{K}_0}h\| = \lim_n \|U^n T^{*n} h\| = \lim_n \|T^{*n}h\| \geqq c\|h\| \qquad (h \in \mathfrak{H})$$

et par conséquent $P_{\mathfrak{K}_0}\mathfrak{H}$ est un ensemble fermé. Vu aussi (1. 4) on aura donc

$$P_{\mathfrak{K}_0}\mathfrak{H} = \mathfrak{K}_0 \quad \text{et} \quad \|P_{\mathfrak{K}_0}h\| \geqq c\|h\| \quad \text{pour tout} \quad h \in \mathfrak{H}. \tag{2. 1}$$

Cela nous permet d'appliquer le lemme de ci-dessus aux décompositions (1. 1) de $\mathfrak{K}$. En vertu de ce lemme, les relations (2. 1) entraînent

$$P^{\mathfrak{L}_*}M_+(\mathfrak{L}) = M_+(\mathfrak{L}_*) \quad \text{et} \quad \|P^{\mathfrak{L}_*}l\| \geqq c\|l\| \quad \text{pour tout} \quad l \in M_+(\mathfrak{L}). \tag{2. 2}$$

Puisque $P^{\mathfrak{L}_*}$ permute à U, les relations (2. 2) subsistent aussi pour $U^{-\nu}M_+$, au lieu de M_+ ($\nu = 0, 1, \ldots$), et comme ces sous-espaces vont en croissant et déterminent l'espace M en question, on aura, par raison de continuité,

$$P^{\mathfrak{L}_*}M(\mathfrak{L}) = M(\mathfrak{L}_*) \quad \text{et} \quad \|P^{\mathfrak{L}_*}l\| \geqq c\|l\| \quad \text{pour tout} \quad l \in M(\mathfrak{L}). \tag{2. 3}$$

Posons

$$Q = P^{\mathfrak{L}_*}|M(\mathfrak{L}), \quad W = U|M(\mathfrak{L}), \quad W_* = U|M(\mathfrak{L}_*). \tag{2. 4}$$

Nous avons, en vertu de (2. 2) et (2. 3),

$$\|Q\| \leqq 1, \quad QM(\mathfrak{L}) = M(\mathfrak{L}_*), \quad QM_+(\mathfrak{L}) = M_+(\mathfrak{L}_*), \quad QW = W_*Q,$$

$$\|Q^{-1}\| \leqq \frac{1}{c}, \quad Q^{-1}M(\mathfrak{L}_*) = M(\mathfrak{L}), \quad Q^{-1}M_+(\mathfrak{L}_*) = M_+(\mathfrak{L}), \quad Q^{-1}W_* = WQ^{-1}.$$

Dans l'hypothèse additionnelle que l'espace $\mathfrak{H}$ (et par conséquent l'espace $\mathfrak{K}$) sont *séparables*, ces relations nous permettent d'appliquer le lemme du n° 2 de [IX] et nous obtenons qu'il existe dans $D_0 = \{\lambda: |\lambda| < 1\}$ deux fonctions analytiques bornées

$$\{\mathfrak{L}, \mathfrak{L}_*, \Theta(\lambda)\} \quad \text{et} \quad \{\mathfrak{L}_*, \mathfrak{L}, \Omega(\lambda)\}.$$

telles que

$$\|\Theta(\lambda)\| \leqq 1 \quad \text{et} \quad \|\Omega(\lambda)\| \leqq \frac{1}{c},$$

et que, en considérant les représentations unitaires

$$\Phi^{\mathfrak{L}} \sum_{-\infty}^{\infty} W^n l_n = \sum_{-\infty}^{\infty} e^{int} l_n \qquad (l_n \in \mathfrak{L}),$$

$$\Phi^{\mathfrak{L}_*} \sum_{-\infty}^{\infty} W_*^n l_n = \sum_{-\infty}^{\infty} e^{int} l_n \qquad (l_n \in \mathfrak{L}_*)$$

de $M(\mathfrak{L})$ sur $L_0^2(\mathfrak{L})$ et de $M(\mathfrak{L}_*)$ sur $L_0^2(\mathfrak{L}_*)$,[2]) on a

$$\Phi^{\mathfrak{L}_*}Qf = \Theta(e^{it})\Phi^{\mathfrak{L}}f \quad \text{pour} \quad f \in M(\mathfrak{L}), \tag{2. 5}$$

---

[2]) Il s'agit des espaces des fonctions $l(t)$ $(0 \leqq t \leqq 2\pi)$, à valeurs vecteurs dans $\mathfrak{L}$ ou $\mathfrak{L}_*$, selon les cas, mesurables et telles que $\|l(t)\|^2$ est intégrable. Le signe $_0$ veut indiquer que l'intégration doit être prise par rapport à la mesure normée $dt/(2\pi)$.

et

(2.6) $$\Phi^{\mathfrak{L}} Q^{-1} g = \Omega(e^{it}) \Phi^{\mathfrak{L}_*} g \quad \text{pour} \quad g \in M(\mathfrak{L}_*).$$ [3]

Il s'ensuit que

$$\Theta(e^{it}) \Omega(e^{it}) \Phi^{\mathfrak{L}_*} g = \Phi^{\mathfrak{L}_*} g \quad \text{pour} \quad g \in M(\mathfrak{L}_*)$$

et

$$\Omega(e^{it}) \Theta(e^{it}) \Phi^{\mathfrak{L}} f = \Phi^{\mathfrak{L}} f \quad \text{pour} \quad f \in M(\mathfrak{L});$$

vu que $\mathfrak{L}$ et $\mathfrak{L}_*$ sont séparables, cela entraîne

$$\Theta(e^{it}) \Omega(e^{it}) = I_{\mathfrak{L}_*}, \quad \Omega(e^{it}) \Theta(e^{it}) = I_{\mathfrak{L}}$$

pour presque tous les $t \in (0,2\pi)$ et par conséquent [4]

$$\Theta(\lambda) \Omega(\lambda) = I_{\mathfrak{L}_*}, \quad \Omega(\lambda) \Theta(\lambda) = I_{\mathfrak{L}} \quad \text{pour tout} \quad \lambda \in D_0,$$

donc $\Omega(\lambda) = \Theta(\lambda)^{-1}$.

Or, la fonction $\{\mathfrak{L}, \mathfrak{L}_*, \Theta(\lambda)\}$ coïncide avec la fonction caractéristique $\{\mathfrak{D}_T, \mathfrak{D}_{T^*}, \Theta_T(\lambda)\}$ de T où

$$\mathfrak{D}_T = \overline{D_T \mathfrak{H}}, \quad \mathfrak{D}_{T^*} = \overline{D_{T^*} \mathfrak{H}}, \quad D_T = (I - T^* T)^{1/2}, \quad D_{T^*} = (I - TT^*)^{1/2},$$

$$\Theta_T(\lambda) = \left[-T + \sum_1^\infty \lambda^n D_{T^*} T^{*n-1} D_T \right] \Big| \mathfrak{D}_T.$$ [5]

Ainsi il résulte que si la contraction T de l'espace (séparable) $\mathfrak{H}$ est similaire à une transformation unitaire, $\Theta_T(\lambda)^{-1}$ existe au sens strict pour tout $\lambda \in D_0$ et est bornée par une constante indépendante de λ.

2. Montrons que la condition que nous venons d'obtenir est aussi suffisante pour que T soit similaire à une transformation unitaire.

En effet, on sait que la relation (2. 5) est vérifiée pour toute contraction T, avec une fonction analytique contractive $\{\mathfrak{L}, \mathfrak{L}_*, \Theta(\lambda)\}$ qui coïncide avec la fonction caractéristique de T; *cf.* [VIII], formule (3.16). Les hypothèses

$$\Theta_T(\lambda) \mathfrak{D}_T = \mathfrak{D}_{T^*}, \quad \|\Theta_T(\lambda)^{-1}\| \leqq \frac{1}{c} \quad (\lambda \in D_0)$$

entraînent

$$\Theta(\lambda) \mathfrak{L} = \mathfrak{L}_*, \quad \|\Theta(\lambda)^{-1}\| \leqq \frac{1}{c} \quad (\lambda \in D_0),$$

donc on aura

(2.7) $$\|P^{\mathfrak{L}_*} f\|_{\mathfrak{K}} = \|\Phi^{\mathfrak{L}_*} P^{\mathfrak{L}_*} f\|_{L_0^2(\mathfrak{L}_*)} = \|\Theta(e^{it}) \Phi^{\mathfrak{L}} f\|_{L_0^2(\mathfrak{L}_*)} \geqq c \|\Phi^{\mathfrak{L}} f\|_{L_0^2(\mathfrak{L})} = c \|f\|_{\mathfrak{K}}$$

pour tout $f \in M(\mathfrak{L})$. De plus, $\{\mathfrak{L}_*, \mathfrak{L}, \Theta(\lambda)^{-1}\}$ étant une fonction analytique bornée, pour toute fonction $u_*(\lambda) \in H_0^2(\mathfrak{L}_*)$ on aura $u(\lambda) = \Theta(\lambda)^{-1} u_*(\lambda) \in H_0^2(\mathfrak{L})$,[6] d'où il

---

[3]) Égalités dans les espaces L_0^2 correspondants, c'est-à-dire pour presque tous les $t \in (0, 2\pi)$.
[4]) *Cf.* [IX], p. 292, note 14.
[5]) *Cf.* [VIII], n° 3, et [IX], p. 292.

s'ensuit que

$$\Theta H_0^2(\mathfrak{L}) \equiv \{\Theta u : u \in H_0^2(\mathfrak{L})\} = H_0^2(\mathfrak{L}_*).$$

Grâce à la relation (2. 5) ce résultat entraîne que

$$P^{\mathfrak{L}_*} M_+(\mathfrak{L}) = M_+(\mathfrak{L}_*). \tag{2. 8}$$

Eu égard aux décompositions (1. 1) de $\mathfrak{K}$, les relations (2. 7) et (2. 8) entraînent, en vertu du lemme du n° 1, que

$$P_{\mathfrak{K}_0}\mathfrak{H} = \mathfrak{K}_0 \quad \text{et} \quad \|P_{\mathfrak{K}_0}h\| \geqq c\|h\| \qquad (h \in \mathfrak{H}). \tag{2. 9}$$

Cela veut dire que la transformation linéaire

$$S = P_{\mathfrak{K}_0}|\mathfrak{H}$$

applique $\mathfrak{H}$ sur $\mathfrak{K}_0$ de manière biunivoque et bicontinue. De la relation (1. 3) nous obtenons que

$$UP_{\mathfrak{K}_0}T^*h = \lim U^{n+1}T^{*n+1}h = P_{\mathfrak{K}_0}h \qquad (h \in \mathfrak{H}),$$

d'où

$$U_0ST^* = S, \quad TS^*U_0^* = S^*, \quad T = S^*U_0S^{*-1},$$

U_0 étant la transformation $U_0 = U|\mathfrak{K}_0$, qui est unitaire dans $\mathfrak{K}_0$.

3. Ainsi, nous avons démontré le suivant

Théorème 1. *Pour que la contraction T d'un espace de Hilbert (séparable) soit similaire à une transformation unitaire, il faut et il suffit que la fonction caractéristique $\{\mathfrak{D}_T, \mathfrak{D}_{T^*}, \Theta_T(\lambda)\}$ satisfasse à la condition que $\Theta_T(\lambda)^{-1}$ existe au sens strict pour tout $\lambda \in D_0$ et est bornée par une constante indépendante de λ. Dans cette condition, T est notamment similaire à la partie $U_0 = U|\mathfrak{K}_0$ de sa dilatation unitaire minimum.*

Ajoutons le fait, démontré dans [VIII], n° 3, que dans le cas où T est complètement non-unitaire, la transformation U_0 est unitairement équivalente à la multiplication par la fonction e^{it} dans l'espace fonctionnel

$$\overline{\Delta_T L_0^2(\mathfrak{D}_T)} \quad \text{où} \quad \Delta_T(t) = [I - \Theta_T(e^{it})^* \Theta_T(e^{it})]^{1/2} \qquad (0 \leqq t \leqq 2\pi). \tag{2. 10}$$

Par conséquent, si la contraction T est similaire à une transformation unitaire, la partie complètement non-unitaire de T est similaire à la multiplication par la fonction e^{it} dans l'espace (2. 10).

3. Une application

1. Envisageons une transformation linéaire dans l'espace (séparable) $\mathfrak{H}$, de la forme suivante:

$$A = R + iQ \tag{3. 1}$$

---

[6]) $H_0^2(\mathfrak{L})$ est le sous-espace de $L_0^2(\mathfrak{L})$, constitué des fonctions $u(e^{it})$, valeurs limites p. p. des fonctions $u(\lambda) = \sum_0^\infty \lambda^n l_n$, analytiques dans D_0, telles que $l_n \in \mathfrak{L}$, $\sum_0^\infty \|l_n\|^2 < \infty$. De même pour $\mathfrak{L}_*$.

où R et Q sont des transformations autoadjointes dont Q est positive et bornée. Cherchons des conditions dans lesquelles A est similaire à une transformation autoadjointe. Il ne restreint évidemment pas la généralité de supposer que $\|Q\|<1$. Dans ce cas $(A\pm iI)^{-1}$ existent au sens strict, conséquence de ce que $(R\pm iI)^{-1}$ existent au sens strict et sont bornées par 1, et de ce que, dans la relation

$$A\pm iI = R\pm iI+iQ = (R\pm iI)[I+i(R\pm iI)^{-1}Q]$$

on a $\|i(R\pm iI)^{-1}Q\| = \|(R\pm iI)^{-1}\|\,\|Q\|<1$. Puisque $Q\geqq O$, la transformée cayleyenne de A,

$$T = (A-iI)(A+iI)^{-1} = I-2i(A+iI)^{-1} \tag{3. 2}$$

est une contraction dans $\mathfrak{H}$. Vu aussi la relation réciproque

$$A = i(I+T)(I-T)^{-1}, \tag{3. 3}$$

il s'ensuit aussitôt que la décomposition de T en ses parties unitaire et complètement non-unitaire, $T=T_1\oplus T_0$, fournit une décomposition $A=A_1\oplus A_0$ en somme orthogonale d'une transformation autoadjointe A_1 et d'une transformation A_0 qui est, dans un sens évident, "complètement non-autoadjointe"[7]), pareille décomposition de A étant déterminée de manière univoque.

Il s'ensuit aussi que si T est similaire à une transformation unitaire, A est similaire (par la même transformation biunivoque et bicontinue S) à une transformation autoadjointe. Pour que ce soit le cas, il faut et il suffit, en vertu du théorème 1, que la fonction caractéristique $\{\mathfrak{D}_T, \mathfrak{D}_{T^*}, \Theta_T(\lambda)\}$ vérifie les conditions

$$\Theta_T(\lambda)\mathfrak{D}_T=\mathfrak{D}_{T^*} \quad \text{pour tout} \quad \lambda\in D_0 \tag{3. 4}$$

et

$$\|\Theta_T(\lambda)g\|\geqq c\|g\| \quad \text{pour tout} \quad g\in\mathfrak{D}_T \text{ et } \lambda\in D_0, \tag{3. 5}$$

avec une constante $c>0$.

Or, dans notre cas, (3. 4) est une conséquence des autres conditions. En effet, puisque $T^{-1}=(A+iI)(A-iI)^{-1}$ existe au sens strict, il s'ensuit de la relation

$$TD_T=D_{T^*}T^{8})$$

que $T\mathfrak{D}_T=\mathfrak{D}_{T^*}$; vu que $\Theta_T(0)=-T|\mathfrak{D}_T$, cela fournit (3. 4) pour $\lambda=0$. Vu aussi (3. 5) pour $\lambda=0$, il résulte que $\Theta_T(0)^{-1}$ existe au sens strict. Donc l'ensemble Λ des points $\lambda\in D_0$ pour lesquels $\Theta_T(\lambda)^{-1}$ existe au sens strict[9]), n'est pas vide. Puisque l'ensemble Λ est évidemment ouvert, on aura démontré que $\Lambda=D_0$ dès qu'on montre que les hypothèses

$$\lambda_n\in\Lambda, \quad \lambda_n\to\lambda_0\in D_0$$

entraînent

$$\lambda_0\in\Lambda.$$

[7]) Ou "simple", dans la terminologie employée dans [1].

[8]) *Cf.* [VIII], note 1.

[9]) Remarquons, sans en faire usage, que Λ est, d'après [VIII] théorème 4, la partie de l'ensemble résolvant de T située dans D_0.

Or, cela s'ensuit de ce que $\|\Theta_T(\lambda_n)^{-1}\| \leqq 1/c$ en vertu de (3. 5), et que dans la relation

$$\Theta_T(\lambda_0) = \Theta_T(\lambda_n)[I+\Theta_T(\lambda_n)^{-1}(\Theta_T(\lambda_0)-\Theta_T(\lambda_n))]$$

on a $\|(\Theta_T(\lambda_0)-\Theta_T(\lambda_n))\| < c$ pour λ_n suffisamment proche de λ_0. Donc nous avons $\Lambda = D_0$ ce qui prouve que (3. 4) est une conséquence de (3. 5).

Vu que $D_T\mathfrak{H}$ est dense dans $\mathfrak{D}_T$, la condition (3. 5) est équivalente à la suivante:

$$\|\Theta_T(\lambda) D_T h\| \geqq c\|D_T h\| \qquad (h \in \mathfrak{H},\ \lambda \in D_0). \tag{3. 6}$$

Calculons ces opérateurs! En posant

$$J = (A+iI)^{-1} \tag{3. 7}$$

on arrive par un calcul simple aux relations

$$D_T^2 = 4J^*QJ, \quad D_{T^*}^2 = 4JQJ^*. \tag{3. 8}$$

Grâce à la relation

$$\Theta_T(\lambda) D_T = D_{T^*}(I-\lambda T^*)^{-1}(\lambda I - T) \quad (\textit{cf.}\ [\text{VIII}],\ \text{n}^\circ\, 2),$$

on a donc

$$\|\Theta_T(\lambda) D_T h\|^2 = 4\big(QJ^*(I-\lambda T^*)^{-1}(\lambda I-T)h,\ \ J^*(I-\lambda T^*)^{-1}(\lambda I - T)h\big).$$

Soit z lié de λ par

$$z = i\frac{1+\lambda}{1-\lambda}; \tag{3. 9}$$

lorsque $|\lambda| < 1$ on a $\operatorname{Im} z > 0$. Nous avons

$$J^*(I-\lambda T^*)(\lambda I - T) =$$

$$= J^*\left[I - \frac{z-i}{z+i}(A^*-iI)^{-1}(A^*+iI)\right]^{-1} \cdot \left[\frac{z-i}{z+i} I - (A-iI)(A+iI)^{-1}\right] =$$

$$= [(z+i)(A^*-iI)-(z-i)(A^*+iI)]^{-1} \cdot [(z-i)(A+iI)-(z+i)(A-iI)]J =$$

$$= (A^*-zI)(zI-A)J.$$

De cette manière,

$$\|\Theta_T(\lambda) D_T h\|^2 = 4\big(Q(A^*-zI)^{-1}(zI-A)Jh, (A^*-zI)^{-1}(zI-A)Jh\big). \tag{3. 10}$$

2. Appliquons ces formules au cas de la transformation

$$Ah(x) = a(x)h(x) + i\int_0^x h(t)\,dt$$

de l'espace $\mathfrak{H} = L^2(0, 1)$ où $a(x)$ est une fonction donnée, mesurable, à valeurs réelles et finies p. p. On a alors $A = R + iQ$ avec

$$Rh(x) = a(x)h(x) + \frac{i}{2}\Big(\int_0^x - \int_x^1\Big) h(t)\,dt, \quad Qh(x) = \frac{1}{2}\int_0^1 h(t)\,dt = \frac{1}{2}(h, e_0)e_0 \tag{3. 11}$$

où $e_0(x) \equiv 1$; donc $Q \geqq O$, $\|Q\| = \frac{1}{2}$.

Le second membre de (3. 10) sera égal à

$$2\left|\left((A^*-zI)^{-1}(zI-A)Jh, e_0\right)\right|^2 = 2\left|\left(Jh, (\bar{z}I-A^*)(A-\bar{z}I)^{-1}e_0\right)\right|^2.$$

Or,

$$(\bar{z}I-A^*)(A-\bar{z}I)^{-1}e_0 = -e_0+2iQ(A-\bar{z}I)^{-1}e_0 =$$
$$= -e_0+i\left((A-\bar{z}I)^{-1}e_0, e_0\right)e_0 = \theta(z)e_0$$

où

(3. 12) $$\theta(z) = -1+i\left((A-\bar{z}I)^{-1}e_0, e_0\right) \qquad (\operatorname{Im} z>0).$$

Par suite (3. 10) prend la forme

(3. 13) $$\|\Theta_T(\lambda)D_T h\|^2 = 2|(Jh, e_0)|^2\cdot|\theta(z)|^2,$$

tandis que (3. 8) donne

(3. 14) $$\|D_T h\|^2 = 4(QJh, Jh) = 2|(Jh, e_0)|^2.$$

Ainsi la condition (3. 6) se réduit à la suivante:

(3. 15) $$|\theta(z)| \geqq c>0 \quad \text{pour} \quad \operatorname{Im} z>0.$$

Calculons $\theta(z)$ de manière explicite! Dans ce but posons

$$u_z = (A+zI)^{-1}e_0 \quad \text{pour} \quad \operatorname{Im} z\neq 0.$$

On a alors

(3. 16) $$\left(a(x)+z\right)u_z(x)+i\int_0^x u_z(t)\,dt = 1 \qquad (0\leqq x\leqq 1),$$

ou, en posant $\left(a(x)+z\right)u_z(x) = v_z(x)$,

$$v_z(x)+i\int_0^x \frac{v_z(t)}{a(t)+z}\,dt = 1.$$

Cette équation a la seule solution

(3. 17) $$v_z(x) = \exp\left(-i\int_0^x \frac{dt}{a(t)+z}\right).$$

En vertu de (3. 12), (3. 16) et (3. 17) on a donc pour $\operatorname{Im} z>0$:

$$\theta(z) = -1+i(u_{-\bar{z}}, e_0) = -v_{-\bar{z}}(1) = -\exp\left(-i\int_0^1 \frac{dt}{a(t)-\bar{z}}\right),$$

ou, en introduisant la fonction de répartition de $a(x)$, c'est-à-dire la fonction

$$\sigma(a) = \operatorname{mes}\{x:\ 0\leqq x\leqq 1,\ \ a(x)\leqq a\} \qquad (-\infty<a<\infty),$$

on a

(3. 18) $$\theta(z) = -\exp\left(-i\int_{-\infty}^{\infty}\frac{d\sigma(a)}{a-z}\right),$$

d'où

$$|\theta(z)| = \exp\left(-\int_{-\infty}^{\infty}\frac{\beta\, d\sigma(a)}{(a-\alpha)^2+\beta^2}\right) \qquad (z=\alpha+i\beta,\ \beta>0). \tag{3.19}$$

Ainsi, la condition (3. 15) est équivalente à la suivante:

$$F(\alpha,\beta) \equiv \int_{-\infty}^{\infty}\frac{\beta}{(a-\alpha)^2+\beta^2}\, d\sigma(a) \leqq M \qquad \left(M = \log\frac{1}{c};\ -\infty<\alpha<\infty,\ \beta>0\right). \tag{3.20}$$

Lorsque (3. 20) est vérifiée, on a pour tout intervalle fini (α_1, α_2) et pour tout $\beta>0$

$$M(\alpha_2-\alpha_1) \geqq \int_{\alpha_1}^{\alpha_2} F(\alpha,\beta)\, d\alpha = \int_{-\infty}^{\infty}\left[\int_{\alpha_1}^{\alpha_2}\frac{\beta\, d\alpha}{(a-\alpha)^2+\beta^2}\right] d\sigma(a) =$$

$$= \int_{-\infty}^{\infty}\left[\operatorname{arc\,tg}\frac{\alpha_2-a}{\beta} - \operatorname{arc\,tg}\frac{\alpha_1-a}{\beta}\right] d\sigma(a).$$

Si α_1, α_2 sont des points de continuité de $\sigma(a)$, il en résulte, en faisant $\beta\to 0$ et en appliquant le lemme de Fatou, que

$$M(\alpha_2-\alpha_1) \geqq \pi\int_{\alpha_1}^{\alpha_2} d\sigma(a) = \pi[\sigma(\alpha_2)-\sigma(\alpha_1)],$$

donc $\sigma(a)$ vérifie la condition de Lipschitz avec la constante M/π.

Réciproquement, pour $\sigma(a)$ vérifiant la condition de Lipschitz avec la constante M/π, on a

$$F(\alpha,\beta) \leqq \frac{M}{\pi}\int_{-\infty}^{\infty}\frac{\beta\, da}{(a-\alpha)^2+\beta^2} = \frac{M}{\pi}\left[\operatorname{arc\,tg}\frac{a-\alpha}{\beta}\right]_{a=-\infty}^{a=+\infty} = M,$$

donc (3. 20) subsiste.

De cette façon, nous avons démontré la partie a) du théorème suivant:

Théorème 2. *Soit A la transformation de $\mathfrak{H} = L^2(0,1)$, définie par*

$$Ah(x) = a(x)h(x) + i\int_0^x h(t)\, dt \qquad (0\leqq x\leqq 1)$$

où $a(x)$ est une fonction donnée, mesurable, à valeurs réelles finies p. p.

a) *Pour que A soit similaire à une transformation autoadjointe, il faut et il suffit que $\sigma(a)$, la fonction de répartition de $a(x)$, soit Lipschitzienne.*

b) *Lorsque $\sigma(a)$ est Lipschitzienne, la partie complètement non-autoadjointe de A est similaire à l'opérateur de multiplication par la fonction ξ dans l'espace des* *fonctions* $\varphi(\xi) \in L^2(\Omega)$ *où*

$$\Omega = \{a : \sigma'(a) > 0\}.$$

Reste à démontrer la partie b).

De (3. 13) et (3. 14) il résulte

$$\|\Theta(\lambda)g\| = |\theta(z)| \cdot \|g\|$$

d'abord pour $g \in D_T\mathfrak{H}$, puis par continuité pour tout $g \in \mathfrak{D}_T$. Donc on a pour tout point $e^{it} \neq 1$ où la limite radiale $\Theta(e^{it})$ existe,

(3. 21) $$\|\Theta(e^{it})g\| = \lim_{z \to \xi} |\theta(z)| \cdot \|g\|$$

où ξ est l'image, située sur l'axe réelle, du point $\lambda = e^{it}$ par l'homographie (3. 9), c'est-à-dire

$$\xi = -\operatorname{cotg} \frac{t}{2},$$

la limite au point ξ étant du côté du demi-plan supérieur, non-tangentielle. Or, si la fonction $\sigma(a)$ est Lipschitzienne, on a

(3. 22) $$\lim_{z \to \xi} \int_{-\infty}^{\infty} \frac{\beta \, d\sigma(a)}{(a-\alpha)^2 + \beta^2} = \pi\sigma'(\xi)$$

pour presque tous les $\xi \in (-\infty, \infty)$, limite non-tangentielle à l'axe réelle; *cf.* [2], p. 123. Ainsi, il résulte de (3. 19), (3. 21) et (3. 22) que

$$\|\Theta(e^{it})g\|^2 = \exp(-\pi\sigma'(\xi)) \cdot \|g\| \qquad (g \in \mathfrak{D}_T)$$

pour presque tous les points e^{it} du cercle unité, et par conséquent

$$(\Delta_T^2(t)g, g) = (1 - \exp(-2\pi\sigma'(\xi))) \cdot (g, g) \qquad (g \in \mathfrak{D}_T),$$

pour presque tous les $t \in (0, 2\pi)$. Cela entraîne, pour ces t,

(3. 23) $$\Delta_T(t)g = (1 - \exp(-2\pi\sigma'(\xi)))^{1/2} g \qquad (g \in \mathfrak{D}_T).$$

Observons encore que (3. 8) et (3. 11) entraînent

$$D_T^2 h = 4J^*QJh = 2(Jh, e_0)J^*e_0 \qquad (h \in \mathfrak{H})$$

et que, en vertu de (3. 14) on a

$$\|D_T h_0\|^2 = 2|(e_0, e_0)|^2 = 2 \quad \text{pour} \quad h_0 = (A + iI)e_0.$$

Puisque $\mathfrak{D}_T = \overline{D_T\mathfrak{H}} = \overline{D_T^2\mathfrak{H}}$, il s'ensuit de ces résultats que $\mathfrak{D}_T$ est constitué de tous les multiples numériques de J^*e_0, donc $\dim \mathfrak{D}_T = 1$.

De cette façon, l'espace $\overline{\Delta_T L_0^2(\mathfrak{D}_T)}$ s'identifie à l'espace $\overline{\delta L_0^2(0, 2\pi)}$ où

$$\delta(t) = \eta\left(-\operatorname{cotg} \frac{t}{2}\right), \quad \eta(\xi) = [1 - \exp(-2\pi\sigma'(\xi))]^{1/2}.$$

Ainsi, il s'ensuit que T_0, la partie complètement non-unitaire de T, est similaire à la multiplication par la fonction e^{it} dans l'espace $\overline{\delta L_0^2(0, 2\pi)}$. Par conséquent $A_0 = i(I+T_0)(I-T_0)^{-1}$, la partie complètement non-autoadjointe de A, sera similaire à l'opérateur de multiplication par $i\frac{1+e^{it}}{1-e^{it}} = -\operatorname{cotg}\frac{t}{2}$ dans $\overline{\delta L_0^2(0, 2\pi)}$.

Envisageons la transformation suivante:

$$f(t) \equiv F(e^{it}) \to \frac{1}{\sqrt{\pi}(1-\xi^2)} F\left(\frac{\xi-i}{\xi+i}\right) \qquad (0 \leqq t \leqq 2\pi;\ -\infty < \xi < \infty).$$

Elle applique $L_0^2(0, 2\pi)$ unitairement sur $L^2(-\infty, \infty)$ de sorte que le sous-espace $\overline{\delta L_0^2(0, 2\pi)}$ sera appliqué sur le sous-espace $\overline{\eta L^2(-\infty, \infty)}$ et que de plus à l'opérateur de multiplication par $-\operatorname{cotg} t/2$ dans le premier sous-espace il correspond l'opérateur de multiplication par ξ dans le second sous-espace. Or, l'espace $\overline{\eta L^2(-\infty, \infty)}$ s'identifie de manière évidente à l'espace $L^2(\Omega)$ où

$$\Omega = \{\xi : \eta(\xi) \neq 0\} = \{\xi : \sigma'(\xi) \neq 0\},$$

ce qui achève la démonstration du théorème.

3. Dans le cas où la transformation A elle-même est complètement non-autoadjointe, donc $A = A_0$, notre théorème fournit (dans la condition, bien entendu, que $\sigma(a)$ soit Lipschitzienne) une transformation autoadjointe similaire à A. Tel est le cas par exemple lorsque

$$a(x) \equiv x, \quad \text{donc} \quad \sigma(a) = \begin{cases} 0 & \text{pour} \quad a \leqq 0 \\ a & \text{pour} \quad 0 \leqq a \leqq 1, \\ 1 & \text{pour} \quad 1 \leqq a < \infty. \end{cases}$$

En effet, supposons qu'il existe un sous-espace $\mathfrak{H}'$ de $L^2(0, 1)$, qui réduisse A à une transformation autoadjointe. Puisque A est bornée, on a alors pour tout $h \in \mathfrak{H}'$

$$O = (A' - A'^*)h = (A - A^*)h = 2iQh = i(h, e_0)e_0,$$

donc $(h, e_0) = 0$. Par conséquent, on a aussi

$$(h, A^n e_0) = (A^{n*}h, e_0) = 0 \quad \text{pour} \quad h \in \mathfrak{H}' \text{ et } n = 1, 2, \ldots.$$

Or,

$$Ae_0 = x + ix = (1+i)x,$$

$$A^2 e_0 = (1+i)\left(1+\frac{i}{2}\right)x^2, \ldots, A^n e_0 = (1+i)\left(1+\frac{i}{2}\right)\cdots\left(1+\frac{i}{n}\right)x^n, \ldots,$$

donc $h \perp x^n$ $(n = 0, 1, \ldots)$ ce qui entraîne $h = 0$. Ainsi, $\mathfrak{H}' = \{0\}$, ce qui prouve que A est complètement non-autoadjointe.

Nous obtenons donc que *la transformation*

$$Ah(x) = xh(x) + i\int_0^x h(t)\,dt$$

Über die Approximation im starken Sinne

Von G. ALEXITS und D. KRÁLIK in Budapest

1. Einleitung

Vor einigen Jahren hat einer von uns[1]) das Problem der Approximation im starken Sinne aufgeworfen und teilweise gelöst. Dieses Problem besteht im folgenden: Bezeichne $\|\alpha_{nk}\|$ eine unendliche Zahlenmatrix und $s_n(f, x)$ die n-te Partialsumme der Entwicklung einer Funktion $f(x)$ nach einem in $[a, b]$ definierten Orthogonalsystem $\{\varphi_n(x)\}$. Wissen wir, daß die Mittel

$$t_n(f, x) = \sum_{k=0}^{\infty} \alpha_{nk} s_k(f, x)$$

die Funktion $f(x)$ im gewöhnlichen Sinne mit der Geschwindigkeit $\omega_n \to 0$ approximieren, daß also

$$\left| \sum_{k=0}^{\infty} \alpha_{nk} \{f(x) - s_k(f, x)\} \right| = O(\omega_n)$$

gilt, so wird gefragt, *mit welcher Geschwindigkeit die starken Mittel*

$$\sum_{k=0}^{\infty} |\alpha_{nk}|\, |f(x) - s_k(f, x)|$$

gegen Null, und ob sie überhaupt gegen Null streben? Diese Fragestellung verschärft das bekannte Problem der starken Summierbarkeit.

Der Unterschied zwischen Approximation im gewöhnlichen und im starken Sinne ist wesentlich. Man betrachte z. B. die $(C, 1)$-Summen der Reihe $\sum_{n=0}^{\infty} (-1)^n$, welche bekanntlich zu $\frac{1}{2}$ streben. Da die Partialsummen s_n dieser Reihe den Wert 1 oder 0 haben, wird die Zahl $\frac{1}{2}$ durch die gewöhnlichen $(C, 1)$-Mittel mit der Annäherungsgeschwindigkeit

$$\left| \frac{1}{n+1} \sum_{k=0}^{n} \left(\frac{1}{2} - s_k \right) \right| \leqq \frac{1}{2(n+1)}$$

[1]) G. ALEXITS, Une contribution à la théorie constructive des fonctions, *Acta Sci. Math.*, **19** (1958), 149—157.

approximiert, während für die starken $(C, 1)$-Mittel

$$\frac{1}{n+1}\sum_{k=0}^{n}\left|\frac{1}{2}-s_k\right| = \frac{1}{2}$$

gilt; d. h. die Approximation im gewöhnlichen Sinn ergibt die Annäherungsgeschwindigkeit $O(n^{-1})$, während für die im starken Sinne nur $O(1)$ folgt.

Wir wollen uns im folgenden mit der starken Approximation gewisser Mittel der Fourierreihe und der Orthogonalpolynom-Entwicklungen, wie auch einiger anderer Orthogonalentwicklungen beschäftigen. Um diese alle einheitlich behandeln zu können, bedienen wir uns des Begriffes der *polynomartigen Orthonormalsysteme* $\{\varphi_n(x)\}$, welcher auf die einheitliche Behandlung ähnlicher Fragen schon mit Erfolg angewendet wurde.[2]) Darunter verstehen wir ein Orthonormalsystem, dessen Kern von folgender Beschaffenheit ist:

$$K_n(t, x) = \sum_{k=0}^{n}\varphi_k(t)\varphi_k(x) = \sum_{\mu=1}^{r}F_\mu(t, x)\sum_{i,j=-p}^{p}\gamma_{i,j,\mu}^{(n)}\varphi_{n+i}(t)\varphi_{n+j}(x),$$

wo die Konstanten $|\gamma_{i,j,\mu}^{(n)}|$ eine von n unabhängige gemeinsame Schranke C_1 haben, die natürlichen Zahlen p und r von n nicht abhängen, und die meßbaren Funktionen $F_\mu(t, x)$ die Relation

$$|F_\mu(t, x)| \leqq \frac{C_2}{|t-x|} \qquad (t \neq x)$$

für alle $x \in [a, b]$ mit absoluten Konstanten C_1 und C_2 erfüllen. Die polynomartigen Orthonormalsysteme umfassen das trigonometrische System, die orthogonalen Polynomsysteme, wie auch das Haarsche und das Walshsche Orthonormalsystem.

Wir werden das Problem der starken Approximation in dem Fall untersuchen, wo $s_n(f, x)$ die n-te Partialsumme der Entwicklung einer $L^2_{\varrho(x)}$-integrierbaren Funktion $f(x)$ nach einem polynomartigen Orthonormalsystem bedeutet, und als starke Mittel die Summen

$$\sum_{k=n}^{N_n}\frac{k^{q-1}}{n^q}|f(x)-s_k(f, x)|$$

mit einem $q>0$ gewählt werden. Es wird sich herausstellen, daß diese mit einem $N_n = O(n)$ $(N_n \geqq n)$ gebildeten starken Summen bis auf einen konstanten Faktor dieselbe Annäherungsgeschwindigkeit haben wie die im gewöhnlichen Sinn erreichbare beste Approximation mit dem System $\{\varphi_n(x)\}$. Darunter verstehen wir folgendes: Bezeichne

$$T_n(x) = \sum_{k=0}^{n}a_k\varphi_k(x)$$

eine Linearform höchstens n-ter Ordnung, welche mit den ersten $n+1$ Funktionen des geordneten Systems $\{\varphi_n(x)\}$ gebildet wurde. Die nicht negative Zahl

$$E_n(f) = \inf_{T_n}\sup_{a\leqq x\leqq b}|f(x)-T_n(x)|$$

2) Vgl. G. Alexits, *Convergence problems of orthogonal series* (Oxford—London—New York—Paris, 1961), 295—300.

heißt die mit den aus $\{\varphi_n(x)\}$ gebildeten Linearformen höchstens n-ter Ordnung erreichbare beste Approximation im Intervall $[a, b]$. Unser Hauptresultat ist ausgedrückt im folgenden

Satz 1. *Bezeichne $\{\varphi_n(x)\}$ ein im Intervall $[a, b]$ definiertes, konstantentreues,*[3] *polynomartiges Orthonormalsystem bezüglich der Gewichtsfunktion $\varrho(x) \geqq 0$. Gelten im Teilintervall $[c, d]$ von $[a, b]$ die Beziehungen*

$$\sum_{k=0}^{n} \varphi_k^2(x) = O(n), \qquad 0 \leqq \varrho(x) \leqq K_1,$$

so gilt in jedem inneren Teilintervall von (c, d) die Beziehung

$$\frac{1}{n^q} \sum_{k=n}^{N_n} k^{q-1} |f(x) - s_k(f, x)| \leqq K_2 E_n(f)$$

gleichmäßig, wenn nur $N_n = O(n)$ ist.

Von den Konsequenzen dieses Satzes möchten wir eine hervorheben. Es ist bekannt, daß die beste polynomiale Annäherung $E_n(f)$ für die Klasse der Funktionen, die eine stetige r-te Ableitung $f^{(r)}(x)$ haben $\left(f^{(0)}(x) \equiv f(x)\right)$, von der Größenordnung $n^{-r}\omega\left(\frac{1}{n}, f^{(r)}\right)$ ist, wo

$$\omega(\delta, f^{(r)}) = \sup_{\substack{|x'-x| \leqq \delta \\ x, x' \in [a, b]}} |f^{(r)}(x') - f^{(r)}(x)|,$$

also den Stetigkeitsmodul von $f^{(r)}(x)$ bedeutet. Danach folgt aus Satz 1 unmittelbar der

Satz 2. *Gilt $0 < m \leqq \varrho(x) \leqq M$ für $x \in [c, d]$, und ist $f(x)$ eine r-mal stetig differenzierbare Funktion, so besteht bei $N_n = O(n)$ in jedem inneren Teilintervall von (c, d) die Beziehung*

$$\frac{1}{n^q} \sum_{k=n}^{N_n} k^{q-1} |f(x) - s_k(f, x)| \leqq \frac{K_3 \omega\left(\frac{1}{n}, f^{(r)}\right)}{n^r}$$

gleichmäßig, wenn die $s_k(f, x)$ die Partialsummen der Entwicklung von $f(x)$ nach dem Orthogonalpolynomsystem, welches durch $\varrho(x)$ bestimmt ist, bedeuten.

Ein Vorteil unseres starken Summationsverfahrens besteht darin, daß wir daraus auf die Approximationseigenschaften der starken Riesz-Summation $(R, n^q, 1)$ schließen können. Bezeichne nämlich

$$h_n(f, x, q) = \frac{1}{n^q} \sum_{k=1}^{n} k^{q-1} |f(x) - s_k(f, x)|$$

das n-te starke $(R, n^q, 1)$-Mittel, so folgt, wie wir es sehen werden, aus Satz 2 leicht der

[3]) Das System $\{\varphi_n(x)\}$ heißt konstantentreu, wenn für alle konstanten Funktionen $f(x) \equiv c$ die Beziehung $s_n(f, x) \equiv c$ gilt $(n = 0, 1, \ldots)$.

Satz 3. *Gilt* $0<m\leqq\varrho(x)\leqq M$ *in* $[c, d]$, *und hat* $f\in L^2_{\varrho(x)}$ *eine* r*-te Ableitung, welche der Lipschitzbedingung*

$$\sup_{x, x'\in[a,b]}\frac{|f^{(r)}(x')-f^{(r)}(x)|}{|x'-x|^\alpha}\leqq K_4 \qquad (0<\alpha\leqq 1)$$

genügt, so approximieren die starken $(R, n^q, 1)$*-Mittel der durch die Gewichtsfunktion* $\varrho(x)$ *bestimmten Orthonormalpolynomentwicklung von* $f(x)$ *in jedem inneren Teilintervall von* (c, d) *gleichmäßig mit der Annäherungsgeschwindigkeit*

$$h_n(f, x, q) = O\left(\frac{1}{n^{r+\alpha}}\right),$$

wenn nur $q>r+\alpha$ *ist.*

Für die (trigonometrische) Fourierentwicklung gelten die entsprechenden Behauptungen im ganzen Grundintervall $[0, 2\pi]$.

Wir wollen endlich noch bemerken, daß der Satz 1 auch über das Approximationsvermögen der Partialsummen $s_n(f, x)$ einer Entwicklung nach einem konstantentreuen polynomartigen Orthonormalsystem einen gewissen Aufschluß gibt. Es zeigt sich nämlich, daß die Indizes der „schlecht" approximierenden Partialsummen ziemlich gleichverteilt von der Dichte Null sind, welche Tatsache wir in unserem Satz 4 genauer beschreiben werden.

2. Beweis der Sätze 1—3

Beim Beweis der Sätze 1—3 stützen wir uns auf das

Fundamentallemma: *Sei* $\{\varphi_n(x)\}$ *ein konstantentreues polynomartiges Orthonormalsystem bezüglich der Gewichtsfunktion* $\varrho(x)\geqq 0$ *und* $f(x)$ *im Grundintervall* $[a, b]$ *beschränkt:* $|f(x)|\leqq M$. *Bestehen im Teilintervall* $[c, d]$ *die Voraussetzungen des Satzes 1, so gilt in einem jeden inneren Teilintervall* $[c+h, d-h]$ *mit* $h>0$ *die Ungleichung*

$$\frac{1}{n^q}\sum_{k=n}^{N_n} k^{q-1}|s_k(f, x)|\leqq KM,$$

wobei die Konstante K *von* c, d, h *und* q, *nicht aber von* n *und* x *abhängt.*

Zum Beweis zerlegen wir die obige Summe in drei Teile:

$$\frac{1}{n^q}\sum_{k=n}^{N_n} k^{q-1}|s_k(f, x)| = \frac{1}{n^q}\sum_{k=n}^{N_n} k^{q-1}\left|\int_a^b f(t)\varrho(t)\sum_{m=0}^{k}\varphi_m(t)\varphi_m(x)\,dt\right|\leqq$$

$$\leqq\frac{1}{n^q}\sum_{k=n}^{N_n} k^{q-1}\left|\int_a^c+\int_d^b\right|+\frac{1}{n^q}\sum_{k=n}^{N_n} k^{q-1}\left|\int_c^{x-\frac{1}{n}}+\int_{x+\frac{1}{n}}^{d}\right|+\frac{1}{n^q}\sum_{k=n}^{N_n} k^{q-1}\left|\int_{x-\frac{1}{n}}^{x+\frac{1}{n}}\right|=$$

$$= S_1+S_2+S_3,$$

wo die natürliche Zahl n so groß sein soll, daß die Punkte $x-\dfrac{1}{n}$, $x+\dfrac{1}{n}$ im Intervall

(c, d) liegen. Da das System $\{\varphi_n(x)\}$ polynomartig ist, erhalten wir für S_1:

$$S_1 \leqq \sum_{\mu=1}^{r} \sum_{i,j=-p}^{p} \frac{1}{n^q} \sum_{k=n}^{N_n} k^{q-1} |\gamma_{i,j,\mu}^{(k)}| \, |\varphi_{k+j}(x)| \left| \left(\int_a^c + \int_d^b \right) f(t) F_\mu(t, x) \varphi_{k+i}(t) \varrho(t)\, dt \right|.$$

Sei $\psi(t)$ die folgende Funktion:

$$\psi(t) = \begin{cases} f(t) F_\mu(t, x) & \text{für} \quad t \in [a, c] \cup [d, b], \\ 0 & \text{sonst.} \end{cases}$$

Das letzte Integral bedeutet dann den $k+i$-ten Fourierkoeffizienten b_{k+i} in der Entwicklung der Funktion $\psi(t)$ nach dem System $\{\varphi_n(x)\}$. Nach der Cauchyschen bzw. Besselschen Ungleichung ergibt sich die folgende Abschätzung:

$$S_1 \leqq \sum_{\mu=1}^{r} \sum_{i,j=-p}^{p} \frac{1}{n^q} \left\{ \sum_{k=n}^{N_n} k^{2q-2} |\gamma_{i,j,\mu}^{(k)}|^2 \varphi_{k+j}^2(x) \sum_{k=n}^{N_n} b_{k+i}^2 \right\}^{1/2} \leqq$$

$$\leqq \sum_{\mu=1}^{r} \sum_{i,j=-p}^{p} \frac{C_1}{n^q} \left\{ \sum_{k=n}^{N_n} k^{2q-2} \varphi_{k+j}^2(x) \left(\left(\int_a^c + \int_d^b \right) f^2(t) F_\mu^2(t, x) \varrho(t)\, dt \right) \right\}^{1/2}.$$

Der Wert des letzten Integrals ist wegen $|t-x| \geqq h$ und $f^2(t) \leqq M^2$ nicht größer, als

$$\frac{M^2 C_2^2}{h^2} \int_a^b \varrho(t)\, dt = K_5^2 M^2,$$

und da $N_n \leqq K_6 n$ ist, erhalten wir auf Grund der Voraussetzungen des Satzes 1 für S_1:

$$S_1 \leqq \sum_{\mu=1}^{r} \sum_{j=-p}^{p} C_1 K_5 M \frac{N_n^{q-1}}{n^q} \left\{ \sum_{k=n}^{N_n} \varphi_{k+j}^2(x) \right\}^{\frac{1}{2}} \leqq \sum_{\mu=1}^{r} \sum_{j=-p}^{p} K_7 M n^{-\frac{1}{2}} \leqq K_8 M.$$

Zur Abschätzung des Ausdrucks

$$S_2 \leqq \sum_{\mu=1}^{r} \sum_{i,j=-p}^{p} \frac{1}{n^q} \sum_{k=n}^{N_n} k^{q-1} |\varphi_{k+j}(x)| \, |\gamma_{i,j,\mu}^{(k)}| \left| \left(\int_c^{x-\frac{1}{n}} + \int_{x+\frac{1}{n}}^{d} \right) f(t) F_\mu(t, x) \varphi_{k+i}(t) \varrho(t)\, dt \right|$$

verfahren wir ganz analog. Wir erhalten wie oben:

$$S_2 \leqq \sum_{\mu=1}^{r} \sum_{i,j=-p}^{p} K_9 n^{-q} n^{q-\frac{1}{2}} \left\{ \left(\int_c^{x-\frac{1}{n}} + \int_{x+\frac{1}{n}}^{d} \right) f^2(t) F_\mu^2(t, x) \varrho(t)\, dt \right\}^{1/2} \leqq$$

$$\leqq \sum_{\mu=1}^{r} \sum_{i,j=-p}^{p} K_9 n^{-\frac{1}{2}} M K_1^{1/2} C_2 \left\{ \left(\int_c^{x-\frac{1}{n}} + \int_{x+\frac{1}{n}}^{d} \right) \frac{dt}{(t-x)^2} \right\}^{1/2}.$$

Der Klammerausdruck ist nicht größer, als $(2n)^{1/2}$; so daß sich endlich für S_2 die Beziehung

$$S_2 \leqq K_{10} M$$

ergibt.

Zur Abschätzung von S_3 machen wir von der speziellen Struktur der Kernfunktion $K_n(t, x)$ keinen Gebrauch. Auf Grund unserer Voraussetzungen ergibt sich nämlich unmittelbar

$$S_3 \leqq \frac{1}{n^q} \sum_{k=n}^{N_n} k^{q-1} \left\{ \int\limits_{x-\frac{1}{n}}^{x+\frac{1}{n}} f^2(t)\varrho(t)\,dt \right\}^{1/2} \left\{ \int\limits_a^b \left(\sum_{m=0}^{k} \varphi_m(t)\varphi_m(x) \right)^2 \varrho(t)\,dt \right\}^{1/2} \leqq$$

$$\leqq K_{11} n^{-q} \sum_{k=n}^{N_n} k^{q-1} M K_1^{1/2} n^{-\frac{1}{2}} \left(\sum_{m=0}^{k} \varphi_m^2(x) \right)^{1/2} \leqq$$

$$\leqq K_{12} n^{-q-\frac{1}{2}} M \sum_{k=n}^{N_n} k^{q-\frac{1}{2}} \leqq K_{13} M n^{-q-\frac{1}{2}} n^{q+\frac{1}{2}} = K_{13} M.$$

Durch Addition der Abschätzungen bezüglich S_1, S_2 und S_3 erhalten wir

$$\frac{1}{n^q} \sum_{k=n}^{N_n} k^{q-1} |s_k(f, x)| \leqq (K_8 + K_{10} + K_{13}) M,$$

womit unsere Behauptung mit $K = K_8 + K_{10} + K_{13}$ bewiesen ist.

Der Beweis des Satzes 1 ist nun recht einfach. Sei nämlich $T_n(x) \equiv \sum_{k=0}^{n} a_k \varphi_k(x)$ die aus den ersten $n+1$ Funktionen des Systems $\{\varphi_n(x)\}$ gebildete Linearform, welche die Funktion $f(x)$ im Intervall $[a, b]$ am besten approximiert. Dann haben wir:

$$\frac{1}{n^q} \sum_{k=n}^{N_n} k^{q-1} |s_k(f, x) - f(x)| \leqq$$

$$\leqq \frac{1}{n^q} \sum_{k=n}^{N_n} k^{q-1} \{ |s_k(f, x) - s_k(T_n, x)| + |s_k(T_n, x) - T_n(x)| + |T_n(x) - f(x)| \}.$$

Auf Grund der evidenten Beziehungen

$$s_k(f, x) - s_k(T_n, x) = s_k(f - T_n, x), \quad s_k(T_n, x) \equiv T_n(x) \text{ für } k \geqq n,$$
$$|T_n(x) - f(x)| \leqq E_n(f),$$

ergibt sich dann aus unserem Fundamentallemma die Behauptung des Satzes 1.

Was den Beweis des Satzes 2 anbetrifft, folgt aus der Bedingung $0 < m \leqq \varrho(x) \leqq M$ bekanntlich[4]) die Beziehung $\sum_{k=0}^{n} \varphi_k^2(x) = O(n)$ in jedem inneren Teilintervall von $[c, d]$, daher ist der Satz 2 auf Grund klassischer Approximationssätze eine Folge des Satzes 1.

---

[4]) G. Freud, Über die starke $(C, 1)$-Summierbarkeit orthogonaler Polynomreihen, *Acta Math. Acad. Sci. Hung.*, **3** (1952), 83—88.

Es seien nun die Voraussetzungen des Satzes 3 erfüllt. Schreiben wir dann die Größe $h_n(f, x, q)$ in der Form

$$h_n(f, x, q) = \frac{1}{n^q} \sum_{k=1}^{n} k^{q-1} |f(x) - s_k(f, x)| \leqq \frac{1}{n^q} \sum_{j=0}^{m} 2^{jq} \frac{1}{2^{jq}} \sum_{k=2^j}^{2^{j+1}-1} k^{q-1} |f(x) - s_k(f, x)|,$$

wo die natürliche Zahl m durch die Ungleichung $2^{m-1} \leqq n < 2^m$ definiert ist. Aus dem Satz 2 folgt

$$\frac{1}{2^{jq}} \sum_{k=2^j}^{2^{j+1}-1} k^{q-1} |f(x) - s_k(f, x)| = O\left(\frac{1}{2^{j(r+\alpha)}}\right).$$

Bei Beachtung von $q > r + \alpha$ erhalten wir also für $h_n(f, x, q)$ die Abschätzung

$$h_n(f, x, q) = O(n^{-q}) \sum_{j=0}^{m} 2^{j(q-r-\alpha)} = O(n^{-q}) 2^{m(q-r-\alpha)} = O(n^{-q}) n^{q-r-\alpha} = O(n^{-r-\alpha}),$$

womit auch der Satz 3 bewiesen ist.

3. Das Approximationsvermögen der Partialsummen

Unsere Sätze 1—3 legen es nahe, daß schon die einzelnen Partialsummen $s_n(f, x)$ im allgemeinen die Approximationsgeschwindigkeit der starken Mittel erreichen, d. h. daß die Partialsummen $s_{n_k}(f, x)$, welche die Funktion $f(x)$ schlechter approximieren als die starken Mittel, nur selten vorkommen können. Die genaue Formulierung dieser Tatsache befindet sich im folgenden

Satz 4. *Sei $f(x)$ eine r-mal stetig differenzierbare Funktion mit $f^{(r)} \in \mathrm{Lip}\, \alpha$ $(0 < \alpha \leqq 1)$. Bezeichne $\lambda(x) > 0$ eine beliebige monoton gegen Unendlich strebende Funktion und $\nu(n)$ bzw. $\bar{\nu}(n)$ die Anzahl jener Indizes $n_k \leqq n$ bzw. $n \leqq n_k \leqq 2n$, für welche im Punkt $x \in (c, d)$ die Ungleichung*

$$|s_{n_k}(f, x) - f(x)| \geqq \frac{\lambda(n_k)}{n_k^{r+\alpha}}$$

besteht. Dann gilt

$$\lim_{n\to\infty} \frac{\nu(n)}{n} = 0 \quad \textit{bzw.} \quad \lim_{n\to\infty} \frac{\bar{\nu}(n)}{n} = 0,$$

wobei die Menge $\{n_k\}$ im allgemeinen mit x variiert.

Nach Satz 3 ist nämlich

$$\frac{K_{14}}{n^{r+\alpha}} \geqq \frac{1}{n^q} \sum_{k=0}^{n} k^{q-1} |s_k(f, x) - f(x)| \geqq$$

$$\geqq \frac{1}{n^q} \sum_{n_k \leqq n} n_k^{q-1} |s_{n_k}(f, x) - f(x)| \geqq \frac{1}{n^q} \sum_{n_k \leqq n} \lambda(n_k) n_k^{q-1-r-\alpha}.$$

Wir wählen $q > r + \alpha$ derart, daß $r + \alpha > q - 1$, dann ist $n_k^{q-1-r-\alpha} = n_k^{-\beta}$ mit $\beta > 0$. Da $\lambda(x)$ nach Annahme beliebig langsam wächst, dürfen wir ohne Beschränkung der Allgemeinheit $\{\lambda(n)\}$ so wählen, daß $\{\lambda(n)n^{-\beta}\}$ monoton abnimmt, und daher statt der letzten Summe $\nu(n)\lambda(n)n^{q-1-r-\alpha}$ geschrieben werden darf. Dann ist

$$\frac{K_{14}}{n^{r+\alpha}} \geqq \frac{\nu(n)}{n}\cdot\frac{\lambda(n)}{n^{r+\alpha}},$$

oder mit anderen Worten

$$\frac{\nu(n)}{n} \leqq \frac{K_{14}}{\lambda(n)} = o(1).$$

Der Beweis unserer zweiten Behauptung verläuft ganz auf dieselbe Weise:

$$\frac{K_2}{n^{r+\alpha}} \geqq \frac{1}{n^q} \sum_{n \leqq n_k \leqq 2n} n_k^{q-1}|s_{n_k}(f, x)-f(x)| \geqq \frac{\bar{\nu}(n)}{n} \frac{\lambda(2n)}{n^{r+\alpha}} 2^{q-1-r-\alpha};$$

also gilt

$$\frac{\bar{\nu}(n)}{n} \leqq \frac{2^{r+\alpha}}{2^{q-1}}\cdot\frac{K_2}{\lambda(2n)} = o(1).$$

Der soeben bewiesene Satz drückt jene Eigenschaft der Partialsummenfolge $\{s_n(f, x)\}$ aus, von welcher wir in der Einleitung gesprochen haben, nämlich daß die „schlechten" Indizes n_k in der Menge aller natürlichen Zahlen von der Dichte Null sind $\left(\frac{\nu(n)}{n} \to 0\right)$, ferner daß sie „ziemlich gleichverteilt" sind, also sich nicht an einzelnen Stellen verdichten, da sie ja auch zwischen n und $2n$ nur „selten" auftreten $\left(\frac{\bar{\nu}(n)}{n} \to 0\right)$.

4. Der Annäherungsgrad der Rieszschen Mittel allgemeiner Orthogonalreihen

In einer früheren Arbeit[5]) haben wir einen Approximationssatz bezüglich der starken de la Vallée—Poussinschen Mittel allgemeiner Orthogonalreihen veröffentlicht, welcher einen Satz von TANDORI[6]) verschärft. Wir wollen nun unseren in [5]) bewiesenen Satz weiter verschärfen:

Satz 5. *Sei $\{\varphi_n(x)\}$ ein beliebiges, in $[a, b]$ definiertes Orthogonalsystem und $\lambda(x)$ eine positive, für $x\to\infty$ monoton gegen Unendlich strebende, von unten konkave Funktion. Sei ferner $\Sigma c_n^2\lambda^2(n) < \infty$ und*

$$S_n(x) = \sum_{k=0}^{n} c_k\lambda(k)\varphi_k(x), \quad s_n(x) = \sum_{k=0}^{n} c_k\varphi_k(x).$$

Sind die starken Mittel

$$\frac{1}{n^q}\sum_{k=n}^{N_n} k^{q-1}|S_k(x)| \qquad (N_n = O(n))$$

[5]) G. ALEXITS—D. KRÁLIK, Über die Approximation mit starken de la Vallée—Poussinschen Mitteln, *Acta Math. Acad. Sci. Hung.*, **16** (1965), 43—49.

[6]) K. TANDORI, Über Approximationen mit allgemeinen Orthogonalreihen, *Annales Univ. Sci. Budapest*, **3—4** (1960/61), 351—356.

auf der Menge $E \subset [a, b]$ beschränkt, so gilt für die starken $(R, n^q, 1)$-Mittel der Orthogonalreihe

$$\sum c_n \varphi_n(x)$$

auf E fast überall die Beziehung

$$\frac{1}{n^q} \sum_{k=n}^{N_n} k^{q-1} |s_k(x) - f(x)| = O_x\left(\frac{1}{\lambda(n)}\right),$$

wobei $f(x)$ die als Limes im Mittel bis auf eine Nullmenge eindeutig bestimmte Funktion bedeutet.

Unser Satz folgt leicht aus dem folgenden reihentheoretischen

Hilfssatz. *Die Teilsummen S_n der Reihe $\sum a_n$ sollen der Bedingung*

$$\frac{1}{n^q} \sum_{k=n}^{N_n} k^{q-1} |S_k| \leqq K$$

genügen, bezeichne ferner $\lambda(x)$ die Funktion im Satz 5. Gibt es eine Indexfolge $\{\mu_m\}$ derart, daß S_{μ_m} wie auch

$$s_{\mu_m} = \sum_{k=0}^{\mu_m} \frac{a_k}{\lambda(k)}$$

gegen die Grenzwerte S bzw. s konvergieren, so gilt die Ungleichung

$$\frac{1}{n^q} \sum_{k=n}^{N_n} k^{q-1} |s_k - s| \leqq \frac{C(K, S)}{\lambda(n)}.$$

Der Beweis des Hilfssatzes verläuft analog zu dem des entsprechenden Hilfssatzes in unserer Note[5]), von dessen Beweis wir somit absehen können.

Aus diesem Hilfssatz erhalten wir den Satz 5 leicht. Wegen des Riesz—Fischerschen Satzes konvergieren die Partialsummen $s_n(x)$ bzw. $S_n(x)$ in $L^2[a, b]$ gegen Funktionen $f(x)$ bzw. $f^*(x)$. Wir können dann eine Indexfolge $\{\mu_m\}$ derart auswählen, daß die Partialsummen $s_{\mu_m}(x)$ und $S_{\mu_m}(x)$ auch punktweise fast überall gegen $f(x)$ bzw. $f^*(x)$ konvergieren. Nach unserem Hilfssatz besteht also auf der Menge E fast überall die behauptete Relation

$$\frac{1}{n^q} \sum_{k=n}^{N_n} k^{q-1} |s_k(x) - f(x)| = O_x\left(\frac{1}{\lambda(n)}\right),$$

womit der Satz 5 bewiesen ist.

(Eingegangen am 4. September 1964)

Über eine Verallgemeinerung der Distributivitätsgleichung

Von M. HOSSZÚ in Miskolc (Ungarn)

1. Die Funktionalgleichung

$$F[G(x, y), u] = H[K(x, u), L(y, u)] \tag{1}$$

ist eine Verallgemeinerung der sogenannten Distributivitätsgleichung, wo wir speziell $G = H$, $F = K = L$ haben:

$$F[G(x, y), u] = G[F(x, u), F(y, u)]. \tag{2}$$

Eine weitere Verallgemeinerung ist

$$F_0[G(x_1, x_2, \ldots, x_n), u] = H[F_1(x_1, u), F_2(x_2, u), \ldots, F_n(x_n, u)]. \tag{3}$$

Hier können alle Veränderlichen x_i, u und die Werte der unbekannten Funktionen F_i, G, H aus verschiedenen Mengen gewählt werden. Im Spezialfall, daß G, x_i, bzw. H, F_i Elemente einer Menge Q bzw. Q' sind und u ein Parameter ist, gibt die Gleichung (3) die Definition des Homotopiebegriffes.

Nämlich bilden dann die Abbildungen $x \to F_i(x, u)$ für irgendwelches festes u einen Homotopismus, undzwar zwischen den Gruppoiden, die durch Q bzw. Q' definiert sind und welche die Funktionen G bzw. H als Gruppoidoperationen besitzen [2].

2. Man kann die folgenden Fragen stellen:

1. Was sind die allgemeinsten (umkehrbaren, stetigen, usw.) Lösungen F_i der Funktionalgleichung (3) für gegebene Funktionen G, H?

2. Was sind die Lösungen, wenn — umgekehrt — die F_i gegeben sind?

Vom algebraischen Standpunkt aus bedeutet die erste Frage, daß wir die Homotopismusrelationen zwischen zwei gegebenen Gruppoiden suchen. Die zweite Frage bedeutet, daß wir solche Gruppoiden suchen, die ein gegebenes System von Homotopismusrelationen haben. [1, 3]

Die Umkehrbarkeitsbedingungen als Nebenbedingungen für die Lösungen bedeuten, daß die gesuchten Homotopismusrelationen bzw. Gruppoiden als Isotopismusrelationen bzw. Quasigruppen auftreten. u hat die Rolle, zu zeigen, daß wir ein System der Homotopismen betrachten, die irgendwelche Bedingungen, wie die der Transivtiität, befriedigen [4].

Die Lösung für H ist eine ganz beliebige Funktion und im Falle der gegebenen F_i haben wir für G die Gestalt

$$G(x_1, x_2, \ldots, x_n) = F_0^{-1}\{H[F_1(x_1), F_2(x_2), \ldots, F_n(x_n)]\},$$

wo F_0^{-1} die inverse Funktion von F_0 bedeutet. Folglich beschäftigen wir uns im Späteren nur mit der ersten Frage, d. h. wir suchen die Lösung F_i der Funktionalgleichung für gegebene G, H.

3. Wir zeigen, daß die Verallgemeinerung (3) der Distributivitätsgleichung (2) sich mit Hilfe der eindeutigen Umkehrbarkeitsbedingungen über die Funktionen auf den Spezialfall $F_i = F$, $H = G$ reduzieren läßt.

Zuerst beschränken wir uns auf den Fall, daß die Funktionen G, H umkehrbar sind.

Definition 1. Eine Funktion G heißt *umkehrbar*, oder eine *Quasigruppenoperation*, wenn die Gleichung

$$y = G(a_1, \ldots, a_{k-1}, x, a_{k+1}, \ldots, a_n) \tag{4}$$

für irgendein k $(k = 1, 2, \ldots, n)$ und für irgendwelche fixierte Werte von a_i und y, eine eindeutige Lösung x hat.

Wir fixieren ein System von Elementen $a_1, a_2, \ldots, a_n$ und bezeichnen dann die Umkehrfunktionen mit $x = G_k(y)$. Offenbar befriedigen die Funktionen G_k die Zusammenhänge

$$\begin{aligned} G[a_1, \ldots, a_{k-1}, G_k(t), a_{k+1}, \ldots, a_n] &= \\ = G_k[G(a_1, \ldots, a_{k-1}, t, a_{k+1}, \ldots, a_n)] &= t \qquad (k = 1, 2, \ldots, n). \end{aligned} \tag{5}$$

Definition 2. G_0 heißt eine *Isotope* von G, wenn die Bedingung

$$\varphi_0[G_0(x_1, x_2, \ldots, x_n)] = G[\varphi_1(x_1), \varphi_2(x_2), \ldots, \varphi_n(x_n)]$$

für irgendwelche eindeutig umkehrbare Abbildungen $x_i \to \varphi_i(x_i)$ gültig ist. G_0 ist eine *Hauptisotope* von G, wenn φ_0 die identische Abbildung ist.

Der Begriff des Isotopismus ist eine Verallgemeinerung des Begriffes des Isomorphismus, wo wir speziell $\varphi_i = \varphi$ $(i = 0, 1, \ldots, n)$ haben.

Es ist offensichtlich, daß jede Isotope einer Funktion G isomorph zu einer Hauptisotopen ist. Es sei nämlich $\bar{G}$ eine Isotope der G, d. h.

$$\bar{G}(x_1, x_2, \ldots, x_n) = \varphi_0^{-1}\{G[\varphi_1(x_1), \varphi_2(x_2), \ldots, \varphi_n(x_n)]\},$$

wo φ_0^{-1} die inverse Abbildung von φ_0 ist. Dann ist $\bar{G}$ isomorph zu der Hauptisotopen

$$G_0(x_1, x_2, \ldots, x_n) = G\{\varphi_1[\varphi_0^{-1}(x_1)], \varphi_2[\varphi_0^{-1}(x_2)], \ldots, \varphi_n[\varphi_0^{-1}(x_n)]\}$$

der G, und die Abbildung $x \to \varphi_0(x)$ erzeugt den Isomorphismus:

$$\varphi_0[\bar{G}(x_1, x_2, \ldots, x_n)] = G_0[\varphi_0(x_1), \varphi_0(x_2), \ldots, \varphi_0(x_n)].$$

Im Falle, daß die Abbildungen φ_i nicht notwendig umkehrbar sind, sprechen wir über Homotopismus statt Isotopismus bzw. über Homomorphismus statt Isomorphismus. Es führt zu keinem Mißverständnis, wenn wir über den Homotopismus usw. der Funktionen statt des Homotopismus der Quasigruppen sprechen.

Definition 3. Die Quasigruppenoperation G_0 heißt eine *Loop-operation*, wenn ein Einheitselement e existiert, so daß

$$G_0(x_1, x_2, \ldots, x_n) = x_k$$

für alle x_k und $x_i = e$ $(i \neq k)$ gilt.

Hilfssatz. *Jede Loop-hauptisotope einer Quasigruppenoperation G hat die Gestalt*

$$G_0(x_1, x_2, \ldots, x_n) = G[G_1(x_1), G_2(x_2), \ldots, G_n(x_n)], \tag{6}$$

wo die Funktionen G_k die mit Hilfe der Gleichung (5) mit irgendwelchen Elementen a_i definierten Umkehrfunktionen sind.

Beweis. Es sei

$$G_0(x_1, x_2, \ldots, x_n) = G[\varphi_1(x_1), \varphi_2(x_2), \ldots, \varphi_n(x_n)]$$

eine Loop-operation mit Einheitselement e. Dann zeigt die Substitution $x_i = e$ $(i \neq k)$, daß

$$\varphi_k(x_k) = G_k[G_0(e, \ldots, x_k, \ldots, e)] = G_k(x_k)$$

ist, wo G_k eine laut (5) mit $a_i = \varphi_i(e)$ definierte Umkehrfunktion bezeichnet.

Umgekehrt, (5) und (6) zeigen, daß wir

$$G_0[G(x_1, a_2, \ldots, a_n), G(a_1, x_2, a_3, \ldots, a_n), \ldots] = G(x_1, x_2, \ldots, x_n)$$

haben. Mit den Bezeichnungen

$$x_i = a_i \quad (i \neq k), \quad e = G(a_1, a_2, \ldots, a_n),$$

$$t = G(a_1, \ldots, a_{k-1}, x_k, a_{k+1}, \ldots, a_n)$$

haben wir also

$$G_0(e, \ldots, t, e, \ldots, e) = t,$$

d. h. G_0 hat wirklich ein Einheitselement e.

4. Wir beweisen den folgenden

Satz 1. *Es seien G, H gegebene Quasigruppenoperationen. Die notwendige und hinreichende Bedingung für die Lösbarkeit der Funktionalgleichung*

$$F_0[G(x_1, x_2, \ldots, x_n)] = H[F_1(x_1), F_2(x_2), \ldots, F_n(x_n)] \tag{3'}$$

ist, daß H eine solche Loop-isotope habe, die homomorph zu einer Loop-isotope von G ist. Man kann die allgemeinste Lösung für F_i folgendermaßen konstruieren: Mit Hilfe von beliebigen Konstanten a_i, b_i bildet man die Umkehrfunktionen G_k, H_k, welche die Bedingungen (5) und

$$H[b_1, \ldots, b_{k-1}, H_k(t), b_{k+1}, \ldots, b_n] = H_k[H(b_1, \ldots, b_{k-1}, t, b_{k+1}, \ldots, b_n)] = t \tag{5'}$$

befriedigen. Dann konstruiert man die Loop-hauptisotopen (6) und

$$H_0(y_1, y_2, \ldots, y_n) = H[H_1(y_1), H_2(y_2), \ldots, H_n(y_n)]. \tag{6'}$$

Die allgemeinste Lösung F_0 ist ein beliebiger Homomorphismus der G_0 zu H_0:

$$F_0[G_0(x_1, x_2, \ldots, x_n)] = H_0[F_0(x_1), F_0(x_2), \ldots, F_0(x_n)]. \tag{7}$$

Ferner, die allgemeinsten Lösungen für F_i $(i > 0)$ mit den Anfangsbedingungen

$$F_i(a_i) = b_i \qquad (i = 1, 2, \ldots, n) \tag{8}$$

sind

(9) $$F_i(x)=H_i\{F_0[G_i^{-1}(x)]\},$$

wo G_i^{-1} *die inverse Funktion der* G_i *bezeichnet, d. h.*

(10) $$G_i^{-1}(x)=G(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n).$$

Beweis. Fixieren wir $x_i=a_i$ $(i\neq k)$ in (3'), so bekommen wir wegen (10) und (8)

$$F_0[G_k^{-1}(x_k)]=H[F_1(a_1), \ldots, F_{k-1}(a_{k-1}), F_k(x_k), F_{k+1}(a_{k+1}), \ldots]= \\ =H[b_1, \ldots, b_{k-1}, F_k(x_k), b_{k+1}, \ldots],$$

folglich haben wir (9). Damit (9) die Funktionalgleichung (3') befriedigt, ist es notwendig und hinreichend, daß

$$F_0[G(x_1, x_2, \ldots, x_n)]=H\big(H_1\{F_0[G_1^{-1}(x_1)]\}, \ldots\big)=H_0\{F_0[G_1^{-1}(x_1)], \ldots\}$$

gilt. Das ist aber eben (7), wenn wir $G_k(x_k)$ statt x_k schreiben und (6) beachten. Endlich, da alle Loop-isotopen der G bzw. H zu einer Hauptisotopen von der Gestalt (6) bzw. (6') mit geeigneten Konstanten a_i, b_i isomorph sind (siehe den Hilfssatz), haben wir die oben ausgesprochene Bedingung für die Lösbarkeit.

5. Bis jetzt reduzierten wir die Funktionalgleichung (3) zu dem Spezialfall $F_i=F$, d. h. zu (7). Wenn (7) eine eindeutig umkehrbare Lösung hat, dann können wir eine weitere Reduktion machen.

Satz 2. *Es sei* $E(x)$ *eine eindeutig umkehrbare partikuläre Lösung der Funktionalgleichung* (7) *für* F_0. *Dann ist*

(11) $$F_0(x)=E[F_G(x)]=F_H[E(x)]$$

die allgemeinste Lösung von (7), wo F_G *die allgemeinste Lösung der Funktionalgleichung*

(12) $$F_G[G(x_1, x_2, \ldots, x_n)]=G[F_G(x_1), F_G(x_2), \ldots, F_G(x_n)]$$

und F_H *die allgemeinste Lösung der analogen Gleichung (mit* H *statt* G*) ist.*

Beweis. Setzen wir (11) in (7) ein, so bekommen wir z. B.

$$E\{F_G[G(x_1, x_2, \ldots, x_n)]\}=H\{E[F_G(x_1)], \ldots\}=E\{G[F_G(x_1), \ldots]\},$$

woraus (12) folgt. Wir bemerken, daß ein ähnliches Ergebnis gilt, wenn die partikuläre Lösung E nur eine einseitige Inverse hat.

Literatur

[1] J. Aczél, *Vorlesungen über Funktionalgleichungen und ihre Anwendungen* (Basel—Stuttgart, 1960).
[2] R. H. Bruck, A survey of binary systems, *Ergebnisse d. Math.*, **20** (Berlin—Göttingen—Heidelberg, 1958).
[3] M. Hosszú, A generalization of the functional equation of distributivity, *Acta Sci. Math.*, **20** (1959), 67—80.
[4] M. Hosszú, Functional equations and algebraic methods in the theory of geometric objects, *Publicationes Math. Debrecen*, **5** (1958), 294—329.

(Eingegangen am 3. Februar 1964)

General theory of summability. I

By RICHARD JAJTE in Łódź (Poland)

In this paper we give the general definition of a summability method of functions defined on a σ-bicompact space. It seems that this definition includes all the cases so far considered, in particular the broad classes of summability methods of functions considered by K. KNOPP and his pupils. Owing to the generality of our methods we may give, on the one hand, a uniform theory and on the other we can find more applications of this theory. The present paper includes the fundamental theorems concerning the most important properties of the methods considered. These theorems constitute a generalization of the well-known theorems for matrix-transformations (the theorems of TOEPLITZ, KNOPP, MAZUR—ORLICZ, HENSTOCK, STEINHAUS; see references at the end of the paper). In a subsequent paper we shall consider some special classes of methods and some applications of the general theory in other branches.

§ 1

Definition 1. A locally bicompact space is called a σ-bicompact space if it is the sum of a sequence of bicompact sets.[1]) Let X be a σ-bicompact Hausdorff space[2]). By $\Omega(X)$ we denote the set of real functions defined and continuous on X with bicompact supports and by $\Omega^+(X)$ the set of functions continuous and non-negative on X with bicompact supports. Let $J(f)$ be a distributive functional defined on $\Omega(X)$ and non-negative on $\Omega^+(X)$. Let μ and $\int f(x)\,d\mu$ denote the Lebesgue measure and integral generated by the functional $J(f)$.[3])

Definition 2. A function $f(x)$ defined on X shall be called convergent in ∞ to the number ξ if for any $\varepsilon>0$ the set

$$\{x: |f(x)-\xi| \geqq \varepsilon\}$$

is contained in a bicompact set.

[1]) A topological space is called a locally bicompact space if every point $x\in X$ possesses a neighbourhood with a bicompact closure and if the whole space X is not bicompact.

[2]) The space X is considered as fixed in all the further considerations.

[3]) See [7], § 6.

Definition 3. Let $S=\{S_\tau\}$ $(0\leqq\tau<\infty)$ be a family of bicompact subsets of the space X such that $S_{\tau'}\subseteqq S_{\tau''}$ if $\tau'<\tau''$ and $\bigcup_{0\leqq\tau<\infty} S_\tau=X$. The improper integral

$$(S)\int f(x)\,d\mu$$

exists and is equal to α if $\int\limits_{S_\tau} f(x)\,d\mu$ exists for $0\leqq\tau<\infty$ and $\lim\limits_{\tau\to\infty}\int\limits_{S_\tau} f(x)\,d\mu=\alpha$.

Definition 4. Let $\Phi=\{\Phi(t,x)\}$ $(t_0\leqq t<T\leqq\infty)$ denote a family of continuous functions defined on X. A function $f(x)$ defined on X is said to be summable by the method $M=M(\mu,S,\Phi)$ to the number ξ if

(a) the integrals $(S)\int\Phi(t,x)f(x)\,d\mu$ exist for every $t\in[t_0,T)$, and

(b) the limit $\lim\limits_{t\to T-}(S)\int\Phi(t,x)f(x)\,d\mu=\xi$ exists.

Evidently, by specifying the space X and the classes S and Φ we can obtain some well-known classes of summability methods for number sequences or for functions defined on the half line (see e. g. [3]). The case considered here is, however, much more general for it includes even the summability of functions defined in non-metric spaces. Let us observe that the term "summability of functions in ∞" has a formal character, since X can stand e.g. for a closed circle its center excluded. Then the center z_0 of this circle shall play the part of ∞ and we obtain a summability method of a function at the point z_0.

Definition 5. A function defined on X is said to be locally bounded if for every point $x_0\in X$ there exists a neighbourhood $U(x_0)$ in which this function is bounded.

For what follows the following conditions are important.

$$(\mathrm{c}_1)\qquad \overline{\lim_{t\to T-}}\,(S)\int|\Phi(t,x)|\,d\mu<\infty\,;$$

$$(\mathrm{c}_2)\qquad \lim_{t\to T-}(S)\int\Phi(t,x)\,d\mu=\alpha\,;$$

$$(\mathrm{c}_2')\qquad \lim_{t\to T-}(S)\int\Phi(t,x)\,d\mu=1\,;$$

(c_3) $\lim\limits_{t\to T-}\int\limits_A\Phi(t,x)\,d\mu=\alpha(A)$ exists for every measurable set A contained in some bicompact set;

(c_3') $\lim\limits_{t\to T-}\int\limits_A\Phi(t,x)\,d\mu=0$ for every measurable set A contained in some bicompact set.

Definition 6. We call a method M a convergence preserving method (a convergence preserving method for null functions) if it sums all μ-measurable locally bounded and in ∞ convergent (convergent to zero) functions. We call a method M permanent (permanent for null functions) if it sums all the measurable, locally bounded and in ∞ convergent (convergent to zero) functions to their ordinary limits.

Theorem 1. *A method M is convergence preserving for null functions if and only if the conditions* (c_1) *and* (c_3) *are satisfied.*

Proof. The necessity of condition (c_3) is evident. To prove the necessity of condition (c_1) we complete the space X with the point x_∞ into the space X_∞. The set of functions continuous on X and having a limit in ∞ equal to zero (we denote it by $C_0(X_\infty)$) is a Banach space with the norm

$$\|f\|_\infty = \sup_{x\in X} |f(x)|.$$

The functional

$$A_{t,\tau}(f) = \int\limits_{S_\tau} \Phi(t, x) f(x)\, d\mu$$

is linear (i.e. additive and continuous) in $C_0(X_\infty)$. We shall show that its norm is equal to

$$\int\limits_{S_\tau} |\Phi(t, x)|\, d\mu$$

Since the set S_τ is bicompact, we have $\mu(S_\tau)=K<\infty$ and there exists for $\eta>0$ an open set $U\supset S_\tau$ with a bicompact closure such that $\mu(U-S_\tau)<\eta$. We denote by A_1 and A_2 the subsets of the set S_τ in which $\Phi(t, x)\geqq\varepsilon/K$, $\Phi(t, x)\leqq-\varepsilon/K$, respectively. The sets A_i, in view of the continuity of the function $\Phi(t,x)$, are bicompact. In virtue of the well-known theorem of Urysohn (see e. g. [7], p. 34) there exists a continuous function $u(x)$ defined on the whole space X and satisfying the conditions

(i) $0\leqq u(x)\leqq 1$,

(ii) $u(x)=1$ for $x\in A_1$ and $u(x)=0$ for $x\in A_2$.

By the same theorem of Urysohn there exists a continuous function $v(x)$ defined on the whole space X and satisfying the conditions

(j) $0\leqq v(x)\leqq 1$,

(jj) $v(x)=1$ for $x\in S_\tau$ and $v(x)=0$ for $x\in X_\infty-U$.

Let us put $\psi(x) = [2u(x)-1]\cdot v(x)$. We have $|\psi(x)|\leqq 1$ and

$$\psi(x) = \begin{cases} 1 & \text{for} \quad x\in A_1, \\ -1 & \text{for} \quad x\in A_2, \\ 0 & \text{for} \quad x\notin U. \end{cases}$$

Obviously $\psi\in C_0(X_\infty)$. Moreover we have

$$A_{t,\tau}(\psi)\geqq \int\limits_{S_\tau} |\Phi(t, x)|\, d\mu - 2\varepsilon.$$

Since $\|\psi\|_\infty=1$ this implies

$$\|A_{t,\tau}\| = \int\limits_{S_\tau} |\Phi(t, x)|\, d\mu.$$

The functional

$$A_t(f) = \lim_{\tau \to \infty} A_{t,\tau}(f)$$

is defined in $C_0(X_\infty)$, therefore according to the Banach—Steinhaus theorem ([1], Theorem 5, p. 80)

$$\sup_{0 \leq \tau < \infty} \|A_{t,\tau}\| = (S) \int |\Phi(t, x)| \, d\mu < \infty.$$

Choose a set S_τ and an open set $U \supset S_\tau$ such that

$$\int_{U - S_\tau} |\Phi(t, x)| \, d\mu < \varepsilon.$$

We have

$$A_t(\psi) = A_{t,\tau}(\psi) + \int_{U - S_\tau} \Phi(t, x) \psi(x) \, d\mu \geqq \|A_{t,\tau}\| - 2\varepsilon + \int_{U - S_\tau} \Phi(t, x) \psi(x) \, d\mu.$$

But

$$\int_{U - S_\tau} \Phi(t, x) \psi(x) \, d\mu > - \int_{U - S_\tau} |\Phi(t, x)| \, d\mu > -\varepsilon.$$

Therefore

$$A_t(\psi) \geqq \|A_{t,\tau}\| - 3\varepsilon.$$

Passing to the limit as $\varepsilon \to 0$ and $\tau \to \infty$ we obtain immediately

$$\|A_t\| = (S) \int |\Phi(t, x)| \, d\mu.$$

Since by assumption $\lim_{t \to T-} A_t(f)$ exists for every $f \in C_0(X_\infty)$ so by the Banach—Steinhaus theorem we have

$$\overline{\lim_{t \to T-}} \|A_t\| < \infty,$$

which ends the proof of the necessity of condition (c_1).

Now we shall prove that the conditions (c_1) and (c_3) are sufficient.

Let $f(x)$ be a measurable and locally bounded function and $\lim_{x \to \infty} f(x) = 0$. We may assume (considering, if necessary, the function $Cf(x)$) that $|f(x)| \leqq 1$. We put

$$\delta(Z) = \lim_{t \to T-} \int_Z \Phi(t, x) \, d\mu.$$

By condition (c_3) this limit exists for any measurable set contained in a bicompact set. Let ε_n be a sequence decreasing to zero. Then there exists an increasing sequence of bicompact sets $\{V_n\}$ such that for $x \notin V_n$ we have $|f(x)| < \varepsilon_n$. Let

$$Z_k^{(n)} = \{x \in V_n \colon k/2^n \leqq f(x) < k + 1/2^n\} \quad \text{for} \quad k = -2^n, -2^n + 1, \ldots, 2^n - 2,$$

$$Z_{2^n - 1}^{(n)} = \{x \in V_n \colon 1 - 1/2^n \leqq f(x) \leqq 1\}.$$

Put

$$\delta_n(f) = 2^{-n} \sum_{k = -2^n}^{2^n - 1} k \delta(Z_k^{(n)}).$$

Now we shall show that the limit $\delta=\lim_{n\to\infty}\delta_n(f)$ exists. Let $m>n$. We have

$$|\delta_m(f)-\delta_n(f)| = \left|2^{-m}\sum_{k=-2^m}^{2^m-1} k\delta(Z_k^{(m)})-2^{-n}\sum_{k=-2^n}^{2^n-1} k\delta(Z_k^{(n)})\right| \leqq$$

$$\leqq\left|\sum_{k=-2^m}^{2^m-1} k/2^m\delta(V_n\cap Z_k^{(m)})-\sum_{k=-2^n}^{2^n-1} k/2^n\delta(Z_k^{(n)})\right|+N\cdot\varepsilon_n,$$

where

$$N=\sup_{t_0\leqq t<T}(S)\int|\Phi(t,x)|\,d\mu.\,^{4)}$$

We observe that the sets $V_n\cap Z_k^{(m)}$ are formed by dividing the sets $Z_k^{(n)}$, thus the sum under the modulus-sign on the right hand side of the inequality can be written as follows

$$\sum_{k=-2^m}^{2^m-1}\delta(V_n\cap Z_k^{(m)})[k/2^m-\alpha_k] \quad\text{with}\quad |\alpha_k-k/2^m|<2^{-n}.$$

Hence

$$|\delta_m(f)-\delta_n(f)|\leqq N\cdot\varepsilon_n+2^{-n}\cdot\sum_{k=-2^m}^{2^m-1}|\delta(V_n\cap Z_k^{(m)})|\leqq N(\varepsilon_n+2^{-n})\to 0$$

for $n\to\infty$, which implies convergence of the sequence $\delta_n(f)$. We shall prove that $M\text{-}\lim_{x\to\infty} f(x)=\delta$. Let $\eta>0$ be given. Putting

$$g_n(x)=\begin{cases}k/2^n & \text{for } x\in Z_k^{(n)},\\ 0 & \text{for } x\notin V_n,\end{cases}$$

we have

$$\left|(S)\int\Phi(t,x)f(x)\,d\mu-\delta\right|\leqq\left|\int_{V_n}\Phi(t,x)[f(x)-g_n(x)]\,d\mu\right|+$$

$$+\left|\int_{V_n}\Phi(t,x)g_n(x)\,d\mu-\delta\right|+\left|(S)\int_{X-V_n}\Phi(t,x)f(x)\,d\mu\right|.$$

We fix n such that $\max(\varepsilon_n N, N2^{-n}, |\delta_n(f)-\delta|)<\eta/4$. Then we have

$$\left|(S)\int\Phi(t,x)f(x)\,d\mu-\delta\right|\leqq N2^{-n}+\left|\int_{V_n}\Phi(t,x)g_n(x)\,d\mu-\delta_n(f)\right|+$$

$$+|\delta_n(f)-\delta|+N\varepsilon_n<3/4\eta+\left|\sum_{k=-2^n}^{2^n-1}k/2^n\int_{Z_k^{(n)}}\Phi(t,x)\,d\mu-\delta_n(f)\right|<\eta$$

for t sufficiently large, which ends the proof of our theorem.

Theorem 2. *A method M is convergence preserving if and only if the conditions* (c_1), (c_2) *and* (c_3) *are satisfied.*

---

4) Without diminishing the generality of our considerations we may assume that $\sup_{0\leqq t<T}(S)\int|\Phi(t,x)|\,d\mu<\infty$.

Proof. The necessity of conditions (c_2) and (c_3) is evident, while necessity of the condition (c_1) follows from the foregoing theorem. The sufficiency of the conditions (c_1), (c_2) and (c_3) follows from theorem 1 by considering the function $g(x) = f(x) - \lim_{x\to\infty} f(x)$.

Theorem 3. *A method M is permanent for null functions if and only if the conditions* (c_1) *and* (c_3') *are satisfied.*

Proof. The necessity of condition (c_3') is evident. Necessity of condition (c_1) follows from theorem 1. The sufficiency follows from the proof of theorem 1 and from the remark that $\delta_n(f)\equiv 0$.

Theorem 4. *A method M is permanent if and only if the conditions* (c_1), (c_2') *and* (c_3') *are satisfied.*

Proof. The sufficiency of conditions (c_1), (c_2') and (c_3') follows from the theorem 3 by considering the function $g(x) = f(x) - \lim_{x\to\infty} f(x)$.

Definition 7. A method M is called a row-finite method if $\Phi(t, x)=0$ beyond a bicompact set Z_t for $t_0 \leqq t < T$.

Theorem 5. *For each convergence preserving method M there exists a row-finite method M_1 equivalent to M for bounded functions (i. e. summing the same measurable and bounded functions to the same limit).*

We omit the easy proof of this theorem.

§ 2

A number ξ is called the limit value of the function $f(x)$ in ∞ if for any bicompact set $Z \subset X$ and any $\varepsilon > 0$ there exists a point $x \in X - Z$ such that $|f(x) - \xi| < \varepsilon$. We say that $+\infty$ $(-\infty)$ is the limit value of the function $f(x)$ if for any bicompact set Z and for any number K there exists a point $x \in X - Z$ such that $f(x) > K$ $(f(x) < K)$. The upper (lower) bound of the limit values shall be denoted by $\overline{\lim} f(x)$ $(\underline{\lim} f(x))$. The interval

$$K(f) = [\underline{\lim}\, f(x), \overline{\lim}\, f(x)]$$

(finite or infinite) shall be called the core of the function $f(x)$ in ∞.

Let $f(x)$ be a measurable, locally bounded function and $M = M(\mu, S, \Phi)$ a permanent method. We set $\varphi(t) = (S)\int \Phi(t, x) f(x) d\mu$ and

$$K_M(f) = [\underline{\lim}_{t\to T-} \varphi(t), \overline{\lim}_{t\to T-} \varphi(t)].$$

It can easily be shown that if $\Phi(t, x) \geqq 0$ then

$$K_M(f) \subseteqq K(f).$$

We call a number ξ the essential limit value in ∞ of the function $f(x)$ if for any $\varepsilon > 0$ and for any bicompact set Z there exists a point $x \in X - Z$ at which the function $f(x)$ is continuous and such that $|f(x) - \xi| < \varepsilon$. We say that $+\infty$ $(-\infty)$

is the essential limit value of the function $f(x)$ if for any bicompact set Z and for any number L there exsits a point $x \in X-Z$ at which the function $f(x)$ is continuous and such that $f(x)>L(f(x)<L)$. The upper (lower) bound of the essential limit values shall be denoted by $\overline{\lim}\operatorname{ess} f(x)$ ($\underline{\lim}\operatorname{ess} f(x)$). The interval

$$K_{\mathrm{ess}}(f)=[\underline{\lim}\operatorname{ess} f(x), \overline{\lim}\operatorname{ess} f(x)]$$

shall be called the essential core of the function $f(x)$ in ∞. We have obviously

$$K_{\mathrm{ess}}(f) \subseteqq K(f).$$

Theorem 6. *Given a fixed measure μ and a family S determining the improper integral $(S)\int \psi(x)\,d\mu$, for each number ξ from the essential core of the function $f(x)$ there exists a method $M(\mu, S, \Phi)$ permanent and positive (i. e. $\Phi(t, x) \geqq 0$) such that $M\text{-}\lim\limits_{x\to\infty} f(x)=\xi$.*

We omit the proof. (See e. g. [2], p. 77).

§ 3

Definition 8. A method $M=M(\mu, S, \Phi)$ is said to satisfy the condition (w) if there exists a family of bicompact sets $\{Z_t\}$ ($t_0 \leqq t < T$) such that $Z_{t'} \subseteqq Z_{t''}$ for $t'<t''$, $\bigcup\limits_{t_0\leqq t<T} Z_t = X$ and $\lim\limits_{t\to T-}(S)\int\limits_{X-Z_t} |\Phi(t, x)|\,d\mu=0$.

Theorem 7. *Let two permanent methods M_1 and M_2 satisfying the condition (w) be given. If M_2 is more general than M_1 for bounded functions (i. e. any measurable bounded function summable by M_1 is also summable by M_2), then this methods are consistent for bounded functions (i. e. any measurable bounded function summable by M_1 to the number ξ is summable by M_2 to the same number).*[5]

Proof. Let $M_1=M_1(\mu_1, S_1, \Phi_1)$, $M_2=M_2(\mu_2, S_2, \Phi_2)$. Suppose, there exists a measurable bounded function $f(x)$ defined on the space X and summable in ∞ by the methods M_i to different numbers. We may assume that $M_1\text{-}\lim f(x)=0$ and $M_2\text{-}\lim f(x)=1$. We shall prove that in this case there exists a measurable and bounded function summable by M_1 but not summable by M_2.

Let $\{Z_t^{(1)}\}$ and $\{Z_t^{(2)}\}$ be two families of bicompact subsets of the space X such that

$$\lim_{t\to T-}(S_i)\int\limits_{X-Z_t^{(i)}} |\Phi_i(t, x)|\,d\mu_i = 0 \qquad (i=1, 2).$$

Put

$$V_t = Z_t^{(1)} \cup Z_t^{(2)}.$$

By W_0 we denote the empty set. Let $W_1=V_{t_0}$. In virtue of the permanency of the methods M_i there exists a number $t_1 \in [t_0, T)$ such that

$$\Big|\int\limits_{W_1} \Phi_i(t, x) f(x)\,d\mu_i\Big| < 1/2 \quad \text{for} \quad t \geqq t_1 \qquad (i=1, 2).$$

---

[5]) Compare [5] and [8].

Put $W_2 = V_{t_1}$. Let us choose $t_2 > \max(t_1, T-\frac{1}{2})$ such that for $t \geqq t_2$

$$\Big|\int\limits_{W_1} \Phi_i(t,x) f(x)\, d\mu_i\Big| < 1/4, \quad \Big|\int\limits_{W_2 - W_1} \Phi_i(t,x) f(x)\, d\mu_i\Big| < 1/4 \qquad (i=1,2).$$

Let $W_3 = V_{t_2}$ etc. We define inductively a number sequence $\{t_k\}$ increasing to T and a sequence of bicompact sets $\{W_k\}$ such that for $t \geqq t_k$ we have

$$\Big|\int\limits_{W_j - W_{j-1}} \Phi_i(t,x) f(x)\, d\mu_i\Big| < 2^{-k} \quad \text{for} \quad j=1,2,\dots,k;\ i=1,2,$$

and $W_k = V_{t_{k-1}}$. Let $\{\xi_k\}$ be a numerical sequence such that $\overline{\lim\limits_{k\to\infty}}\, \xi_k = 1$, $\underline{\lim\limits_{k\to\infty}}\, \xi_k = 0$, $|\xi_k| \leqq 1$, $\lim\limits_{k\to\infty} |\xi_k - \xi_{k-1}| = 0$. Put $g(x) = \xi_k f(x)$ for $x \in W_k - W_{k-1}$ $(k=1,2,\dots)$. Then we have for $i=1,2$

$$(S_i)\int \Phi_i(t,x) g(x)\, d\mu_i = \sum_{j=1}^{k} \xi_j \int\limits_{W_j - W_{j-1}} \Phi_i(t,x) f(x)\, d\mu_i +$$

$$+ \xi_{k+1} \int\limits_{W_{k+2} - W_k} \Phi_i(t,x) f(x)\, d\mu_i + (\xi_{k+2} - \xi_{k+1}) \int\limits_{W_{k+2} - W_{k+1}} \Phi_i(t,x) f(x)\, d\mu_i +$$

$$+ (S_i) \int\limits_{X - V_{t_{k+1}}} \Phi_i(t,x) g(x)\, d\mu_i.$$

The following estimations hold for the terms on the right hand side of our equality:

$$\Big|(\xi_{k+2} - \xi_{k+1}) \int\limits_{W_{k+2} - W_{k+1}} \Phi_i(t,x) f(x)\, d\mu_i\Big| \leqq$$

$$\leqq |\xi_{k+2} - \xi_{k+1}| \sup_{t_0 \leqq t < T} (S)\int |\Phi_i(t,x)|\, d\mu_i \cdot \sup_{x\in X} |f(x)|,$$

$$\Big|\sum_{j=1}^{k} \xi_j \int\limits_{W_j - W_{j-1}} \Phi_i(t,x) f(x)\, d\mu_i\Big| \leqq k\cdot 2^{-k} \quad \text{for} \quad t \geqq t_k,$$

$$\Big|(S_i) \int\limits_{X - V_{t_{k+1}}} \Phi_i(t,x) g(x)\, d\mu_i\Big| \leqq (S_i) \int\limits_{X - Z_t^{(i)}} |\Phi_i(t,x)|\, d\mu_i \cdot \sup_{x\in X} |f(x)| \quad \text{for} \quad t_k \leqq t < t_{k+1},$$

$$\Big|\int\limits_{W_{k+2} - W_k} \Phi_i(t,x) f(x)\, d\mu_i - (S_i)\int \Phi_i(t,x) f(x)\, d\mu_i\Big| \leqq$$

$$\leqq \Big|\int\limits_{W_k} \Phi_i(t,x) f(x)\, d\mu_i\Big| + \Big|(S_i) \int\limits_{X - W_{k+2}} f(x) \Phi_i(t,x)\, d\mu_i\Big| \leqq$$

$$\leqq k\cdot 2^{-k} + (S_i) \int\limits_{X - Z_t^{(i)}} |\Phi_i(t,x)|\, d\mu_i \cdot \sup_{x\in X} |f(x)| \quad \text{for} \quad t_k \leqq t < t_{k+1}.$$

Considering in the above inequality k as a function of t given by the inequality

$t_k \leqq t < t_{k+1}$ we see that

$$\int_{W_{k+2}-W_k} \Phi_i(t, x) f(x)\, d\mu_i \to M_i\text{-}\lim_{x\to\infty} f(x) \qquad (i=1, 2)$$

as $t \to T-$, while the other terms on the right hand side of the inequality tend to zero as $t \to T-$. From the divergence of the sequence $\{\xi_k\}$ we see that the generalized limit M_2-$\lim g(x)$ does not exist while M_1-$\lim g(x) = M_1$-$\lim f(x) = 0$, which ends the proof of our theorem.

Theorem 8. *Let a convergence preserving method $M(\mu, S, \Phi)$ be given. Let $f(x)$ be a measurable, bounded function defined on X and let $|f(x)| \leqq \lambda$. Put*

$$\omega_\alpha(x) = \begin{cases} 1 & \textit{if } f(x) \leqq \alpha, \\ 0 & \textit{if } f(x) > \alpha. \end{cases}$$

If the limit M-$\lim\limits_{x\to\infty} \omega_\alpha(x) = \omega(\alpha)$ exists for any $|\alpha| < \lambda$ then the function $\omega(\alpha)$ has a finite variation in the interval $[-\lambda, \lambda]$, the limit M-$\lim\limits_{x\to\infty} f(x)$ exists, and

$$M\text{-}\lim_{x\to\infty} f(x) = \int_{-\lambda}^{\lambda} \alpha\, d\omega(\alpha).$$

Proof. Put

$$h_t(\alpha) = (S)\int \Phi(t, x)\, \omega_\alpha(x)\, d\mu.$$

Let us observe that for any t the functions $h_t(\alpha)$ have bounded variations with the common bound $\sup\limits_{t_0 \leqq t < T} (S)\int |\Phi(t, x)|\, d\mu$. This follows from the estimation

$$\sum_{i=1}^{n-1} |h_t(\alpha_{i+1}) - h_t(\alpha_i)| \leqq (S)\int |\Phi(t, x)| \sum_{i=1}^{n-1} \left(\omega_{\alpha_{i+1}}(x) - \omega_{\alpha_i}(x)\right) d\mu \leqq$$

$$\leqq (S)\int |\Phi(t, x)|\, d\mu, \quad \text{where} \quad -\lambda \leqq \alpha_1 < \alpha_2 < \ldots < \alpha_n \leqq \lambda.$$

If for any α the limit

$$\lim_{t\to T-} h_t(\alpha) = \omega(\alpha)$$

exists, then obviously $w(\alpha)$ has a finite variation and

$$\operatorname{Var} \omega(\alpha) \leqq \sup_{t_0 \leqq t < T} (S)\int |\Phi(t, x)|\, d\mu.$$

Put $Z_i = \{x: \omega_{\alpha_{i+1}}(x) > \omega_{\alpha_i}(x)\}$ and denote by $\chi_Z(x)$ the characteristic function of the set Z. Let $|\alpha_{i+1} - \alpha_i| < \varepsilon$ $(i = 1, 2, \ldots, n-1)$. We have

$$\left|(S)\int \Phi(t, x) f(x)\, d\mu - \sum_{i=1}^{n-1} \alpha_i [h_t(\alpha_{i+1}) - h_t(\alpha_i)]\right| =$$

$$= \left|(S)\int \Phi(t, x) \left[f(x) - \sum_{i=1}^{n-1} \alpha_i \chi_{Z_i}(x)\right] d\mu\right| \leqq$$

$$\leqq \sup_{x\in X} |f(x) - \sum \alpha_i \chi_{Z_i}(x)| \cdot (S)\int |\Phi(t, x)|\, d\mu \leqq \varepsilon \cdot (S)\int |\Phi(t, x)|\, d\mu.$$

Hence it follows that

$$(S)\int \Phi(t, x)f(x)\,d\mu = \int_{-\lambda}^{\lambda} \alpha\, dh_t(\alpha).$$

Since

$$\left|\int_{-\lambda}^{\lambda} \alpha\, d[\omega(\alpha)-h_t(\alpha)]\right| \leqq [\alpha|\omega(\alpha)-h_t(\alpha)|]_{-\lambda}^{\lambda} + \int_{-\lambda}^{\lambda} |\omega(\alpha)-h_t(\alpha)|\, d\alpha \to 0 \text{ as } t\to T-,$$

then

$$\lim_{t\to T-} (S)\int \Phi(t, x)f(x)\,d\mu = \int_{-\lambda}^{\lambda} \alpha\, d\omega(\alpha),$$

which ends the proof.

Theorem 9. *For each permanent method M there exists a measurable function taking only the values* 0 *and* 1, *which is not summable by the method M.*

Proof. Suppose that there exists a permanent method summing all measurable functions with the values 0 and 1. By the foregoing theorem this method would sum all measurable and bounded functions. Let $\{t_n\}$ be a sequence increasing to T. The method $\tilde{M}$ described by the functional

$$\tilde{M}(f) = \lim_{n\to\infty} (S)\int \Phi(t_n, x)f(x)\,d\mu$$

would then sum all the measurable and bounded functions and it would satisfy the condition (w) (definition 8). Thus there would exist a method satisfying the condition (w) and not weaker for bounded functions than all the permanent methods. Thus by the consistence theorem (theorem 7) all permanent methods satisfying the condition (w) would be consistent for bounded functions, which is impossible. This contradiction proves our theorem.

References

[1] S. Banach, *Théorie des opérations linéaires* (Warszawa, 1932).
[2] R. G. Cooke, *Infinite matrices and sequence spaces* (London, 1950).
[3] G. H. Hardy, *Divergent series.* (Oxford, 1956).
[4] R. Jajte, On a theorem of Toeplitz, *Colloquium Math.* (To appear)
[5] S. Mazur and W. Orlicz, Sur les méthodes de sommation, *Comptes Rendus Acad. Sci. Paris*, **196** (1933), 32—34.
[6] S. Mazur and W. Orlicz, On linear methods of summability, *Studia Math.*, **14** (1954), 129—160.
[7] M. A. Наймарк, *Нормированные кольца* (Moscow, 1956).
[8] G. Petersen, Summability methods and bounded sequences, *J. London Math. Soc.*, **31** (1956), 324—326.
[9] H. Steinhaus, Remarks on the generalization of the idea of limit (in Polish), *Prace Mat.-Fiz.* **22** (1911), 121—134.
[10] O. Toeplitz, Über allgemeine lineare Mittelbildungen, *Prace Mat.-Fiz.*, **22** (1911), 113—120.

UNIVERSITY ŁÓDŹ

(Received November 13, 1963)

Über verschiedene Konvergenzarten der trigonometrischen Reihen. II (Die Entwicklung der Ableitungen)

Von LÁSZLÓ LEINDLER in Szeged

Einleitung

Sei $f(x)$ eine 2π-periodische, auf $(0, 2\pi)$ quadratisch integrierbare Funktion mit der Fourier-Entwicklung

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \equiv S[f]. \tag{1}$$

Bezeichne $E_n = E_n^{(2)}(f)$ den besten Annäherungsgrad von $f(x)$, im Sinne der Metrik von $L^2(0, 2\pi)$, mit trigonometrischen Polynomen $(n-1)$-ter Ordnung; nach dem Satz von GRAM [2] ist

$$E_n^2 = \pi \sum_{k=n}^{\infty} (a_k^2 + b_k^2) = \pi \sum_{k=n}^{\infty} \varrho_k^2.$$

Wenn wir im Folgenden über die Existenz der p-ten Ableitung einer Funktion $f(x)$ sprechen (p eine natürliche Zahl), so verstehen wir, dass $f(x)$ fast überall gleich der p-fach iterierten Integralfunktion einer integrierbaren Funktion $g(x)$ ist, und dann nennen wir $g(x)$ die p-te Ableitung von $f(x)$, in Formel: $g(x) = f^{(p)}(x)$.

Satz I. *Unter der Bedingung*

$$\int_0^1 \frac{1}{t^{p+\frac{3}{2}}} \left(\int_0^{2\pi} [f(x+2t) + f(x-2t) - 2f(x)]^2 \, dx \right)^{1/2} dt < \infty \tag{2}$$

gilt

$$\sum_{k=1}^{\infty} k^p (|a_k| + |b_k|) < \infty.$$

Satz II. *Unter der Bedingung*

$$\int_0^1 \frac{|\log t|}{t^{p+1}} \left(\int_0^{2\pi} [f(x+2t) + f(x-2t) - 2f(x)]^2 \, dx \right)^{1/2} dt < \infty \tag{3}$$

existiert $f^{(p)}(x)$ und konvergiert ihre Fourier-Entwicklung fast überall unbedingt, d. h. bei jeder Anordnung ihrer Glieder.

Satz III. *Unter der Bedingung*

$$(4)\qquad \int_0^1 \frac{(|\log t|)^{1/2}}{t^{p+1}} \left(\int_0^{2\pi} [f(x+2t)+f(x-2t)-2f(x)]^2\,dx \right)^{1/2} dt < \infty$$

existiert $f^{(p)}(x)$ *und konvergieren die Reihen*

$$\frac{A_0}{2} + \sum_{n=1}^{\infty} \pm (A_n \cos nx + B_n \sin nx) \qquad \left(\begin{matrix} A_n \\ B_n \end{matrix} = \frac{1}{\pi} \int_0^{2\pi} f^{(p)}(x) \begin{matrix} \cos nx \\ \sin nx \end{matrix} dx \right)$$

bei fast allen Vorzeichenverteilungen gleichmäßig.

Satz IV. *Ist*

$$(5)\qquad \int_0^1 \int_0^{2\pi} \frac{|\log t|}{t^{2p+1}} (f(x+2t)+f(x-2t)-2f(x))^2\,dx\,dt < \infty,$$

so existiert $f^{(p)}(x)$ *und konvergiert ihre Fourier-Entwicklung fast überall.*

Satz V. *Unter der Bedingung*

$$(6)\qquad \int_0^1 \int_0^{2\pi} \frac{1}{t^{2p+2}} (f(x+2t)+f(x-2t)-2f(x))^2\,dx\,dt < \infty$$

existiert $f^{(p)}(x)$ *und wenn sie eine* 2π*-periodische, stetige Funktion ist, so ist die Fourier-Entwicklung von* $f^{(p)}(x)$ *gleichmäßig konvergent.*

Man könnte noch weitere ähnliche Sätze, z. B. für die Summation und für verschiedene Arten von Approximationen und Summationen (siehe die entsprechenden Sätze mit $p=0$ in [4], [5]) leicht beweisen.

Wir setzen

$$w_2^{(2)}(\delta, f) = \left\{ \frac{1}{\delta} \int_0^{\delta} \left(\int_{-\pi}^{\pi} [f(x+2t)+f(x-2t)-2f(x)]^2\,dx \right) dt \right\}^{1/2}.$$

Satz VI. *Unter der Bedingung*

$$\sum_{n=1}^{\infty} n^{p-\frac{1}{2}} w_2^{(2)}\left(\frac{1}{n}, f\right) < \infty$$

existiert $f^{(p)}(x)$ *und hat eine absolut konvergente Fourier-Entwicklung.*

Satz VII. *Ist*

$$\sum_{n=2}^{\infty} n^{p-1} (\log n)\, w_2^{(2)}\left(\frac{1}{n}, f\right) < \infty,$$

so existiert $f^{(p)}(x)$ *und konvergiert ihre Fourier-Entwicklung fast überall unbedingt.*

§ 1. Hilfssätze

Hilfssatz I. *Ist $f(x) \in L^2(0, 2\pi)$, so gilt*

$$E_n \leqq w_2^{(2)}\left(\frac{1}{n}, f\right) \qquad \text{(s. [4], Hilfssatz II).}$$

Hilfssatz II. *Es sei $\lambda(x)$ eine positive, monoton nichtwachsende Funktion mit $\lambda^{-1}(n) = O(n^\beta)$ $(\beta > 0)$. Dann folgt aus der Bedingung*

$$\text{(1. 1)} \qquad \int_0^1 \frac{1}{t^2 \lambda\left(\frac{1}{t}\right)} \left(\int_0^{2\pi} [f(x+2t)+f(x-2t)-2f(x)]^2\, dx\right)^{1/2} dt < \infty$$

die Ungleichung

$$\text{(1. 2)} \qquad \sum_{n=1}^{\infty} \frac{1}{\lambda(n)} E_n < \infty$$

und aus der Bedingung

$$\text{(1. 3)} \qquad \int_0^1 \frac{1}{t^2 \lambda\left(\frac{1}{t}\right)} \int_0^{2\pi} [f(x+2t)+f(x-2t)-2f(x)]^2\, dx\, dt < \infty$$

die Ungleichung

$$\text{(1. 4)} \qquad \sum_{n=1}^{\infty} \frac{1}{\lambda(n)} E_n^2 < \infty.$$

Beweis. Die Implikation (1. 3)$\Rightarrow$(1. 4) haben wir schon in [5] (s. den Hilfssatz III) bewiesen. Die Implikation (1. 1)$\Rightarrow$(1. 2) beweisen wir ähnlich zum Beweis des Satzes I von [4]. Ist $f(x)$ in $(0, 2\pi)$ fast überall konstant, so ist die Behauptung klar. Wir nehmen an, daß $f(x)$ nicht fast überall konstant ist. Nach der Parsevalschen Gleichung gilt

$$\text{(1. 5)} \qquad F(t) = \int_0^{2\pi} [f(x+2t)+f(x-2t)-2f(x)]^2\, dx = 16\pi \sum_{k=1}^{\infty} \varrho_k^2 \sin^4 kt.$$

Nach den Obigen ist die Funktion

$$\alpha(t) = \left(t^2 \lambda\left(\frac{1}{t}\right) \sqrt{F(t)}\right)^{-1}$$

fast überall endlich. Aus der Bedingung (1. 1) und aus der Definition von $\alpha(t)$ ergibt sich, daß

$$\text{(1. 6)} \qquad \int_0^1 \frac{1}{t^4 \lambda^2\left(\frac{1}{t}\right) \alpha(t)}\, dt < \infty$$

und

(1. 7)
$$\int_0^1 \int_0^{2\pi} \alpha(t)[f(x+2t)+f(x-2t)-2f(x)]^2\,dx\,dt<\infty$$

sind. Aus (1. 5) und (1. 7) folgt, auf Grund des Satzes von FUBINI,

(1. 8)
$$\sum_{k=1}^{\infty} \varrho_k^2 \int_0^1 \alpha(t)\sin^4 kt\,dt<\infty.$$

Wir bezeichnen mit I_k die Menge derjenigen Punkte t von $[0, 1]$, für die $\sin^4 kt \geqq \frac{1}{4}$ ist. Bezeichne

$$I(k)=\int_{I_k} \alpha(t)\,dt.$$

Nach (1. 8) besteht

(1. 9)
$$\sum_{k=1}^{\infty} \varrho_k^2 I(k)<\infty.$$

Es sei $\{p_m\}$ $(p_0=0)$ eine Folge von natürlichen Zahlen, für die

(1. 10)
$$2^{m\beta} \leqq \sum_{n=p_{m-1}}^{p_m-1} \frac{1}{\lambda(n)} \leqq K_1 2^{m\beta} \qquad (K_1 \geqq 2)$$

$(m=1, 2, \ldots)$ erfüllt ist. Die Existenz einer solchen Folge $\{p_m\}$ ist auf Grund der bezüglich $\lambda(x)$ getroffenen Bedingungen klar. Wir definieren nun die Folge $\{k_m\}$ folgenderweise: es sei k_m die kleinste unter den natürlichen Zahlen n, für die $p_m \leqq n < p_{m+1}$ und

$$I(n)=\min_{p_m \leqq k < p_{m+1}} \{I(k)\}$$

gelten. Wir zeigen, daß

(1. 11)
$$\sum_{m=1}^{\infty} \frac{2^{2\beta m}}{I(k_m)}<\infty$$

ist. Zuerst beweisen wir, daß

(1. 12)
$$I(k_m) \geqq \frac{1}{2^{8+4\beta}}\, 2^{2\beta m} \left(\int_{\frac{1}{p_{m-2}}}^{\frac{1}{p_{m-3}}} \frac{1}{t^4 \lambda^2\left(\frac{1}{t}\right) \alpha(t)}\,dt \right)^{-1}$$

ist. Es sei $\Delta_m=\left[\frac{1}{p_{m-2}}, \frac{1}{p_{m-3}}\right]$. Durch Anwendung der Cauchyschen Ungleichung erhält man

$$(1.13)\qquad \left(\int\limits_{\Delta_m\cap I_{k_m}} \frac{1}{t^2\lambda\left(\frac{1}{t}\right)}\,dt\right)^2 \leqq \int\limits_{\Delta_m\cap I_{k_m}} \alpha(t)\,dt\cdot \int\limits_{\Delta_m\cap I_{k_m}} \frac{1}{t^4\lambda^2\left(\frac{1}{t}\right)\alpha(t)}\,dt \leqq$$

$$\leqq l(k_m)\int\limits_{\Delta_m} \frac{1}{t^4\lambda^2\left(\frac{1}{t}\right)\alpha(t)}\,dt.$$

Das Integral links kann man wegen der Monotonität von $\lambda(x)$ und $\lambda^{-1}(n)=O(n^\beta)$ folgenderweise abschätzen:

$$(1.14)\qquad \int\limits_{\Delta_m\cap I_{k_m}} \frac{1}{t^2\lambda\left(\frac{1}{t}\right)}\,dt \geqq \frac{1}{2^4}\int\limits_{\Delta_m} \frac{1}{t^2\lambda\left(\frac{1}{t}\right)}\,dt = \frac{1}{2^4}\int\limits_{\bar{\Delta}_m} \frac{1}{\lambda(x)}\,dx,$$

wobei $\bar{\Delta}_m$ das Intervall $[p_{m-3}, p_{m-2}]$ bezeichnet. Nach (1. 10) ist

$$\int\limits_{\bar{\Delta}_m} \frac{1}{\lambda(x)}\,dx \geqq \sum_{n=p_{m-3}}^{p_{m-2}-1} \frac{1}{\lambda(n)} \geqq \frac{1}{2^{2\beta}}\,2^{m\beta}.$$

Daraus und aus (1. 13) und (1. 14) folgt die Ungleichung (1. 12). Die Ungleichung (1. 11) bekommen wir aus (1. 12) auf Grund von (1. 6) sofort. Um den Beweis der Implikation (1. 1)$\Rightarrow$(1. 2) zu beweisen, müssen wir noch zeigen, daß (1. 2) aus (1. 9) und (1. 11) folgt. Durch einfache Rechnung ergibt sich:

$$\frac{1}{\sqrt{\pi}}\sum_{n=1}^{\infty}\frac{1}{\lambda(n)}E_n = \sum_{m=0}^{\infty}\sum_{n=p_m}^{p_{m+1}-1}\frac{1}{\lambda(n)}\left(\sum_{k=n}^{\infty}\varrho_k^2\right)^{1/2} \leqq$$

$$\leqq \sum_{m=0}^{\infty}\sum_{n=p_m}^{p_{m+1}-1}\frac{1}{\lambda(n)}\sum_{\nu=m}^{\infty}\left(\sum_{k=p_\nu}^{p_{\nu+1}-1}\varrho_k^2\right)^{1/2} \leqq$$

$$\leqq K_1\sum_{m=0}^{\infty}2^{m\beta}\sum_{\nu=m}^{\infty}\left(\sum_{k=p_\nu}^{p_{\nu+1}-1}\varrho_k^2\right)^{1/2} \leqq 2^\beta K_1\sum_{m=0}^{\infty}2^{m\beta}\left(\sum_{k=p_m}^{p_{m+1}-1}\varrho_k^2\right)^{1/2}.$$

Daraus bekommen wir weiter durch Anwendung der Cauchyschen Ungleichung

$$\frac{1}{2^\beta K_1\sqrt{\pi}}\sum_{n=1}^{\infty}\frac{1}{\lambda(n)}E_n \leqq \sum_{m=0}^{\infty}\sqrt{l(k_m)}\left(\sum_{k=p_m}^{p_{m+1}-1}\varrho_k^2\right)^{1/2}\cdot\frac{2^{m\beta}}{\sqrt{l(k_m)}} \leqq$$

$$\leqq \left\{\sum_{m=0}^{\infty} l(k_m)\sum_{k=p_m}^{p_{m+1}-1}\varrho_k^2\sum_{m=0}^{\infty}\frac{2^{2\beta m}}{l(k_m)}\right\}^{1/2},$$

woraus das Erfülltsein der Bedingung (1. 2) auf Grund der Definition von $l(k_m)$ und der Ungleichungen (1. 9) und (1. 11) folgt.

Damit haben wir den Hilfssatz II bewiesen.

§ 2. Beweis der Sätze

Beweis von Satz I. Aus der Bedingung (2) auf Grund des Hilfssatzes II mit $\lambda(x)=x^{1/2-p}$ ergibt sich

$$\sum_{n=1}^{\infty} n^{\left(p-\frac{1}{2}\right)} E_n < \infty.$$

Hieraus folgt

$$\sum_{k=1}^{\infty} k^p(|a_k|+|b_k|)<\infty \tag{2.1}$$

wegen

$$\sum_{k=3}^{\infty} k^p(|a_k|+|b_k|) \leqq 2 \sum_{m=1}^{\infty} \left\{ \sum_{k=2^m+1}^{2^{m+1}} k^{2p} \sum_{k=2^m+1}^{2^{m+1}} \varrho_k^2 \right\}^{1/2} \leqq$$

$$\leqq 2^{p+1} \sum_{m=1}^{\infty} 2^{m\left(p+\frac{1}{2}\right)} E_{2^m} \leqq 2^{2p+2} \sum_{m=1}^{\infty} \sum_{n=2^{m-1}+1}^{2^m} n^{p-\frac{1}{2}} E_n < \infty.$$

Beweis von Satz II.

Aus (3) nach dem Hilfssatz II mit $\lambda(x)=\dfrac{x^{1-p}}{|\log x|}$ folgt

$$\sum_{n=2}^{\infty} n^{p-1} (\log n) E_n < \infty.$$

Daraus

$$\sum_{n=3}^{\infty} \frac{1}{n} \sqrt{\sum_{k=n}^{\infty} k^{2p} \varrho_k^2} \leqq 2 \sum_{m=1}^{\infty} \sqrt{\sum_{k=2^m+1}^{\infty} k^{2p} \varrho_k^2} \leqq \tag{2.2}$$

$$\leqq 2 \sum_{m=1}^{\infty} \sqrt{\sum_{\nu=m}^{\infty} 2^{2(\nu+1)p} \sum_{k=2^\nu+1}^{2^{\nu+1}} \varrho_k^2} \leqq 2^{p+1} \sum_{m=1}^{\infty} \sum_{\nu=m}^{\infty} 2^{\nu p} E_{2^\nu} \leqq$$

$$\leqq 2^{p+1} \sum_{\nu=1}^{\infty} \nu 2^{\nu p} E_{2^\nu} \leqq 2^{2p+3} \sum_{\nu=1}^{\infty} \sum_{n=2^{\nu-1}+1}^{2^\nu} n^{p-1} (\log n) E_n =$$

$$= 2^{2p+3} \sum_{n=2}^{\infty} n^{p-1} (\log n) E_n < \infty.$$

Dies ergibt insbesondere

$$\sum_{k=1}^{\infty} \varrho_k^2 k^{2p} < \infty,$$

folglich existiert $f^{(p)}(x)$ und gehört zu $L^2(0,2\pi)$. Somit gilt nach einem bekannten Satz (s. z. B. [8] S. 40): $S^{(p)}[f]=S[f^{(p)}]$. Nach einem Satz des Verfassers [3] konvergiert $S[f^{(p)}]$ fast überall unbedingt, wenn

$$\sum_{n=1}^{\infty} \frac{1}{n} E_n^{(2)}(f^{(p)}) < \infty \tag{2.3}$$

ist. Um den Satz II zu beweisen, genügt es also (2.3) zu zeigen.

Aus (2. 2) folgt aber (2. 3), weil

$$E_n^{(2)}(f^{(p)}) = \left(\sum_{k=n}^{\infty} k^{2p} \varrho_k^2\right)^{1/2}$$

ist.

Beweis von Satz III. Nach einem Satz von SALEM und ZYGMUND [7] genügt es zu zeigen, daß

$$\sum_{n=2}^{\infty} \frac{1}{n\,[\log n]^{1/2}} E_n^{(2)}(f^{(p)}) < \infty \tag{2. 4}$$

ist. Da $S^{(p)}[f] = S[f^{(p)}]$ auch jetzt gilt, so ist $E_n^{(2)}(f^{(p)}) = \left\{\sum_{k=n}^{\infty} k^{2p} \varrho_k^2\right\}^{1/2}$. Nach dem Hilfssatz II mit $\lambda(x) = \frac{x^{1-p}}{[|\log x|]^{1/2}}$ ist

$$\sum_{n=2}^{\infty} n^{p-1} [\log n]^{1/2} E_n < \infty$$

wegen (4). Daraus kann man (2. 4) durch eine einfache, zu (2.2) analoge Rechnung herleiten.

Beweis von Satz IV. Aus (5) folgt

$$\sum_{n=1}^{\infty} n^{2p-1} (\log n) E_n^2 < \infty,$$

woraus wir die Ungleichung

$$\sum_{k=2}^{\infty} k^{2p} \varrho_k^2 \log k < \infty \tag{2. 5}$$

folgenderweise erhalten können:

$$\sum_{k=3}^{\infty} k^{2p} \varrho_k^2 \log k = O(1) \sum_{k=3}^{\infty} k^{2p} \varrho_k^2 \sum_{n=3}^{k} \frac{1}{n} = O(1) \sum_{n=3}^{\infty} \frac{1}{n} \sum_{k=n}^{\infty} k^{2p} \varrho_k^2 =$$

$$= O(1) \sum_{m=1}^{\infty} \sum_{n=2^m+1}^{2^{m+1}} \frac{1}{n} \sum_{\nu=m}^{\infty} \sum_{k=2^\nu+1}^{2^{\nu+1}} k^{2p} \varrho_k^2 = O(1) \sum_{m=1}^{\infty} \sum_{\nu=m}^{\infty} 2^{2p\nu} E_{2^\nu}^2 =$$

$$= O(1) \sum_{\nu=1}^{\infty} \nu 2^{2p\nu} E_{2^\nu}^2 = O(1) \sum_{\nu=1}^{\infty} \sum_{n=2^{\nu-1}+1}^{2^\nu} (\log n) n^{2p-1} E_n^2 < \infty.$$

Aus (2. 5) und $S^{(p)}[f] = S[f^{(p)}]$ folgt die Behauptung des Satzes IV nach dem wohlbekannten Kolmogoroff—Seliverstoff—Plessnerschen Satz.

Beweis von Satz V. Aus (6) ergibt sich nach dem Hilfssatz II mit $\lambda(x) = x^{-2p}$, daß

$$\sum_{n=1}^{\infty} n^{2p} E_n^2 < \infty$$

ist. Daraus folgt

$$\sum_{k=1}^{\infty} k^{2p+1} \varrho_k^2 = O(1) \sum_{n=1}^{\infty} n^{2p} E_n^2 < \infty;$$

Also gibt es eine monotone ins Unendliche strebende $\{\mu_k\}$ derart, daß auch

$$\sum_{k=1}^{\infty} \mu_k k^{2p+1} \varrho_k^2 < \infty$$

besteht. Daraus erhalten wir

$$\sum_{k=1}^{n} k[k^p(|a_k|+|b_k|)] \leqq \left[\sum_{k=1}^{n} k\mu_k[k^p(|a_k|+|b_k|)]^2\right]^{1/2} \left[\sum_{k=1}^{n} \frac{k}{\mu_k}\right]^{1/2} = o(n). \tag{2. 6}$$

Da $S^{(p)}[f]=S[f^{(p)}]$ ist, so folgt aus (2. 6) und aus der Stetigkeit von $f^{(p)}(x)$ nach einem bekannten Satz (s. z. B. [1] S. 177), daß die Fourier-Entwicklung von $f^{(p)}(x)$ gleichmäßig konvergiert, was zu beweisen war.

Auf Grund des Hilfssatzes I und der Beweise der Sätze I und IV kann man die Sätze VI—VII leicht beweisen.

Schriftenverzeichnis

[1] Н. К. Бари, *Тригонометрические ряды* (Москва, 1961).
[2] J. P. Gram, Über die Entwicklung reeller Funktionen in Reihen Mittels der Methode der kleinsten Quadrate, *J. reine angew. Math.*, **94** (1883), 41—73.
[3] L. Leindler, Über unbedingte Konvergenz der Orthogonalreihen mit strukturellen Bedingungen, *Studia Math.*, **23** (1963), 113—117.
[4] L. Leindler, Über verschiedene Konvergenzarten trigonometrischer Reihen, *Acta Sci. Math.*, **25** (1964), 233—249.
[5] L. Leindler, Über hinreichende Strukturbedingungen für Fourierreihen, *Publ. Math. Inst. Hung. Acad. Sci.*, **10** (1965).
[6] A. Plessner, Über Konvergenz von trigonometrischen Reihen, *J. reine angew. Math.*, **15** (1926), 15—25.
[7] R. Salem—A. Zygmund, Some properties of trigonometric series, whose terms have random signs, *Acta Math.*, **91** (1954), 245—301.
[8] A. Zygmund, *Trigonometric series.* I—II (Cambridge, 1959).

(Eingegangen am 27. Januar 1964)

Cesàro operators*)

By ARLEN BROWN, P. R. HALMOS, A. L. SHIELDS in Ann Arbor (Michigan, U. S. A.)

Introduction

If f is a sequence of complex numbers, $f=\langle f(0), f(1), f(2), \ldots\rangle$, the sequence $C_0 f$ of averages plays a role in the theory of Cesàro limits; by definition

$$(C_0 f)(n) = \frac{1}{n+1} \sum_{i=0}^{n} f(i)$$

for $n=0, 1, 2, \ldots$. Our study of Cesàro operators began with the following questions. Is it true that if $f \in l^2$, then $C_0 f \in l^2$? If it is true, is the linear transformation C_0 bounded? If C_0 is bounded, what is its spectrum? Along with these discrete questions, it is natural to ask the corresponding continuous ones; they concern the operator C_1 defined on $L^2(0, 1)$ by

$$(C_1 f) = \frac{1}{x} \int_0^x f(y)\, dy$$

for $0<x<1$, and the operator C_∞ defined on $L^2(0, \infty)$ by

$$(C_\infty f)(x) = \frac{1}{x} \int_0^x f(y)\, dy$$

for $0<x<\infty$.

It turns out that all three Cesàro operators (that is, C_0, C_1, and C_∞) are everywhere defined bounded linear transformations on their respective Hilbert spaces (that is, on l^2, $L^2(0, 1)$, and $L^2(0, \infty)$). For C_0 and C_∞ this fact is proved by HARDY, LITTLEWOOD, and PÓLYA [5, Chapter IX]; the proof below (Theorem 1) is somewhat more conceptual and less computational than theirs.

For C_0 we completely determine the norm, the spectrum, and the various parts of the spectrum (Theorem 2). There is, however, much about C_0 that remains unknown. Thus, for instance, very little is known about the structure of the lattice of invariant subspaces of C_0 — a problem that belongs to a subject of great current

*) Research supported in part by a grant from the National Science Foundation.

interest. Another instance: while we prove that C_0 is hyponormal (Theorem 3), the problem of whether or not it is subnormal remains open.

In view of our incomplete information about C_0, it may be surprising to learn that the structures of C_1 and C_∞ are completely known. We prove that $1-C_1^*$ is a unilateral shift of multiplicity 1 (Theorem 4), and $1-C_\infty^*$ is a bilateral shift of multiplicity 1 (Theorem 5). (The operator C_1 has been studied by DE BRANGES also [3]; our methods are completely different from his.) From these facts, via the Beurling theory [1], it is easy to determine the spectra of C_1 and C_∞, and to derive a satisfactory description of their invariant subspace lattices.

Boundedness

The proof that the Cesaro operators are bounded can be made to depend on a criterion due essentially to I. SCHUR [7]. (In the notation of the statement below, SCHUR discusses the case $p(x)\equiv 1$ only; his proof is different from ours. Cf. also [6, Chapter X].) Since this criterion does not seem to be explicit in the literature, we proceed to state and to prove it with sufficent generality to make it appropriate for most applications.

Schur test. *If X is a measure space, if $k(\geqq 0)$ is a measurable function on $X\times X$, if $p(>0)$ is a measurable function on X, and if α and β are constants such that*

$$\int k(x,y)p(y)\,dy \leqq \alpha p(x)$$

and

$$\int k(x,y)p(x)\,dx \leqq \beta p(y),$$

then the equation

$$(Af)(x)=\int k(x,y)f(y)\,dy$$

defines an operator (a bounded linear transformation) on L^2, and $\|A\|^2\leqq\alpha\beta$.

Proof. If f is a bounded measurable function that vanishes outside some measurable set of finite measure, then

$$\int\left|\int k(x,y)f(y)\,dy\right|^2 dx = \int\left|\left(\int \sqrt{k(x,y)}\sqrt{p(y)}\right)\cdot\left(\frac{\sqrt{k(x,y)}}{\sqrt{p(y)}}f(y)\right)dy\right|^2 dx \leqq$$

$$\leqq \int\left(\int k(x,y)p(y)\,dy\right)\cdot\left(\int\frac{k(x,y)}{p(y)}|f(y)|^2\,dy\right)dx \leqq$$

$$\leqq \int \alpha p(x)\left(\int\frac{k(x,y)}{p(y)}|f(y)|^2\,dy\right)dx =$$

$$=\alpha\int\frac{|f(y)|^2}{p(y)}\left(\int k(x,y)p(x)\,dx\right)dy \leqq \alpha\int\frac{|f(y)|^2}{p(y)}\beta p(y)\,dy = \alpha\beta\|f\|^2.$$

Since the functions such as f are dense in L^2, the proof is complete.

Theorem 1. *Each of the Cesàro operators C_0, C_1, and C_∞ is bounded.*

Proof. For C_0 consider the measure space $\{0, 1, 2, \ldots\}$ with the counting measure, and let the kernel k_0 be defined by

$$k_0(i,j) = \begin{cases} 0 & \text{if } 0 \leqq i < j, \\ \dfrac{1}{i+1} & \text{if } 0 \leqq j \leqq i. \end{cases}$$

If $p_0(n) = \dfrac{1}{\sqrt{n+1}}$, then

$$\sum_j k_0(i,j)p_0(j) = \sum_{j=0}^{i} \frac{1}{i+1}\frac{1}{\sqrt{j+1}} <$$

$$< \frac{1}{i+1}\int_0^i \frac{dx}{\sqrt{x}} = \frac{1}{i+1}2\sqrt{i} < \frac{1}{i+1}2\sqrt{i+1} = 2p_0(i).$$

If $j \neq 0$, then

$$\sum_i k_0(i,j)p_0(i) = \sum_{i=j}^{\infty} \frac{1}{i+1}\frac{1}{\sqrt{i+1}} <$$

$$< \int_{j-1}^{\infty} \frac{dx}{(x+1)^{3/2}} = \frac{2}{\sqrt{j}} = \frac{2}{\sqrt{j+1}}\frac{\sqrt{j+1}}{\sqrt{j}} \leqq 2\sqrt{2}p_0(j).$$

Since also

$$\sum_i k_0(i,0)p_0(i) = 1 + \sum_{i=1}^{\infty} k_0(i,0)p_0(i) < 1+2 = 3p_0(0),$$

it follows that

$$\sum_i k_0(i,j)p_0(i) < 3p_0(j)$$

for all j, and the Schur test implies the boundedness of C_0.

For C_1 the measure space is $(0, 1)$ with Lebesgue measure, and the kernel is defined by

$$k_1(x,y) = \begin{cases} 0 & \text{if } 0 < x \leqq y, \\ \dfrac{1}{x} & \text{if } 0 < y < x. \end{cases}$$

If $p_1(x) = \dfrac{1}{\sqrt{x}}$, then

$$\int_0^1 k_1(x,y)p_1(y)\,dy = \frac{1}{x}\int_0^x \frac{dy}{\sqrt{y}} = \frac{1}{x}2\sqrt{x} = 2p_1(x),$$

and

$$\int_0^1 k_1(x,y)p_1(x)\,dx = \int_y^1 \frac{dx}{x^{3/2}} = \frac{2}{\sqrt{y}} - 2 < 2p_1(y),$$

and the Schur test applies again.

For C_∞ the measure space is $(0, \infty)$ with Lebesgue measure, and the kernel k_∞ is defined formally the same way as k_1; the difference is that x and y now vary in $(0, \infty)$ instead of $(0, 1)$. If, as before, $p_\infty(x) = \frac{1}{\sqrt{x}}$, then

$$\int_0^\infty k_\infty(x, y) p_\infty(y)\, dy = \frac{1}{x} \int_0^x \frac{dy}{\sqrt{y}} = \frac{2}{\sqrt{x}} = 2p_\infty(x),$$

and

$$\int_0^\infty k_\infty(x, y) p_\infty(x)\, dx = \int_y^\infty \frac{dx}{x^{3/2}} = \frac{2}{\sqrt{y}} = 2p_\infty(y),$$

and, once more, the Schur test yields the desired result.

An examination of the proof of Theorem 1 yields (via the last assertion of the Schur test) estimates for the norms of C_0, C_1, and C_∞. For C_0 this estimate turns out to be quite crude, and even for C_1 and C_∞, where it is sharp, the method is not sharp enough to tell what the norms of the operators actually are. To settle this question, and others, we turn now to detailed separate examinations of the three Cesàro operators.

The discrete Cesàro operator

Since C_0 is defined on a sequence space, it is naturally associated with a matrix, which is in fact just the kernel k_0. Since

$$k_0 = \begin{pmatrix} 1 & 0 & 0 & \\ \frac{1}{2} & \frac{1}{2} & 0 & \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \\ & & & \ddots \end{pmatrix}, \quad k_0^* = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \\ 0 & \frac{1}{2} & \frac{1}{3} & \\ 0 & 0 & \frac{1}{3} & \\ & & & \ddots \end{pmatrix},$$

it follows that

$$k_0 k_0^* = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{3} & \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \\ & & & \ddots \end{pmatrix}.$$

It turns out therefore that the product $C_0 C_0^*$ is almost the same as the sum $C_0 + C_0^*$; the difference $C_0 + C_0^* - C_0 C_0^*$ is the diagonal operator D_0 with matrix

$$\begin{pmatrix} 1 & 0 & 0 & \\ 0 & \frac{1}{2} & 0 & \\ 0 & 0 & \frac{1}{3} & \\ & & & \ddots \end{pmatrix}.$$

Since $(1-C_0)(1-C_0^*) = 1-D_0$, it follows that

$$\|1-C_0\| = 1,$$

and hence that $\|C_0\| \leqq 2$.

It is perhaps worth while to remark that there are other ways of proving the last inequality. One way is to compute $C_0C_0^*$ immediately, and then apply the Schur test to it (with the same p_0 as in the proof of Theorem 1). Since $C_0C_0^*$ is Hermitian, only half the computation is necessary, and, moreover, the inequalities do yield the sharp result $\|C_0C_0^*\| \leqq 4$. To infer, via this approach, that C_0 itself is bounded, one more step is necessary; we need to know that if k is an infinite matrix with rows in l^2 such that kk^* is bounded, then k itself is bounded (cf. [7] and [5, Chapter VIII]). The proof of this can be carried out by looking at the n-th section $k^{(n)}$ of k and showing that the n-th section of kk^* domaintes $k^{(n)}k^{(n)*}$. (Recall that an infinite matrix is bounded if and only if its sections are uniformly bounded.)

It is easy to prove that the inequality $\|C_0\| \leqq 2$ cannot be improved:

$$\|C_0\| = 2.$$

Indeed if $f_\alpha(n) = \dfrac{1}{(n+1)^\alpha}\left(\alpha > \dfrac{1}{2},\ n=0, 1, 2, \ldots\right)$, then $f_\alpha \in l^2$ and $\|C_0^* f_\alpha\| \to 2\,\|f_\alpha\|$ as $\alpha \to \frac{1}{2}+$. The proof of the latter assertion is a straightforward computation. Since $(C_0^* f_\alpha)(m) = \sum_{n=m}^{\infty} \dfrac{1}{(n+1)^{\alpha+1}}$, $m=0, 1, 2, \ldots$, it follows that

$$\|C_0^* f_\alpha\|^2 = \sum_{m=0}^{\infty}\left(\sum_{n=m}^{\infty}\frac{1}{(n+1)^{\alpha+1}}\right)^2 > \sum_{m=0}^{\infty}\left(\int_{m+1}^{\infty}\frac{dx}{x^{\alpha+1}}\right)^2 = \sum_{m=0}^{\infty}\left(\frac{1}{\alpha}\,\frac{1}{(m+1)^{\alpha}}\right)^2 =$$

$$= \frac{1}{\alpha^2}\sum_{m=0}^{\infty}\frac{1}{(m+1)^{2\alpha}} = \frac{1}{\alpha^2}\,\|f_\alpha\|^2,$$

and this implies the limit assertion.

For our next purpose we need the following lemma: if A is an operator such that $\|A\| \leqq 1$ and if $\|Af\| = \|f\|$ for some nonzero vector f, then $\|A^*g\| = \|g\|$ for some non-zero vector g. For the proof, write $g = Af$, so that $\|g\| = \|f\|$, and observe that

$$\|f\|^2 = (A^*Af, f) \leqq \|A^*Af\| \cdot \|f\| \leqq \|f\|^2.$$

It follows that $\|A^*Af\| = \|f\|$, so that $\|A^*g\| = \|g\|$.

We know that the supremum of $\|C_0 f\|$ (and hence of $\|C_0^* f\|$) for vectors f on the unit sphere is 2; we shall show that the supremum is not attained. Since $\|(1-D_0)f\| < \|f\|$ unless $f=0$, it follows that

$$\|(1-C_0^*)f\|^2 = ((1-C_0)(1-C_0^*)f, f) \leqq \|(1-C_0)(1-C_0^*)f\| \cdot \|f\| < \|f\|^2$$

unless $f=0$. The preceding paragraph is applicable, and we may infer that both $\|(1-C_0)f\|$ and $\|(1-C_0^*)f\|$ are strictly less than $\|f\|$, except when $f=0$. It follows of course that $\|C_0 f\|$ and $\|C_0^* f\|$ are strictly less than $2\,\|f\|$, except when $f=0$. (Proof: $\|C_0 f\| = \|f-(1-C_0)f\| \leqq \|f\| + \|(1-C_0)f\|$.)

The following statement sums up what we have just proved about norms and what we shall go on to prove about spectra.

Theorem 2. (1) $\|1-C_0\| = 1$ *and* $\|C_0\|=2$. (2) If $\|f\|=1$, *then* $\|(1-C_0)f\|<1$ *and* $\|(1-C_0^*)f\|<1$. (3) *The point spectrum of* C_0 *is empty.* (4) *If* $|1-\lambda|<1$, *then* λ *is a simple proper value of* C_0^*. (5) *The point spectrum of* C_0^* *is the open disc* $\{\lambda: |1-\lambda|<1\}$: (6) *The spectrum of* C_0 *is the closed disc* $\{\lambda: |1-\lambda|\leqq 1\}$.

Proof. (1) and (2) were proved above. To prove (3), observe first that if $C_0 f=g$, then $f(0)=g(0)$, and if $n\geqq 1$, then $f(n) = (n+1)\,g(n)-ng(n-1)$. Consequently, if $C_0 f=\lambda f$, then $f(n) = \lambda\big((n+1)f(n)-nf(n-1)\big)$ or $\big(\lambda(n+1)-1\big)f(n) = \lambda\, nf(n-1)$ whenever $n\geqq 1$. If m is the smallest integer for which $f(m)\neq 0$, then $\lambda = \dfrac{1}{m+1}$, so that $0<\lambda\leqq 1$. It follows that if $n\geqq 1$, then

$$|f(n)| = \left|\frac{\lambda n}{\lambda n-(1-\lambda)} f(n-1)\right| \geqq |f(n-1)|,$$

which, for a non-zero f in l^2, is impossible.

To prove (4), observe first that $(C_0^* f)(n) = \sum_{i=n}^{\infty} \frac{1}{i+1} f(i)$ (cf. the matrix k_0^*). If $C_0^* f=g$, then $f(n) = (n+1)\big(g(n)-g(n+1)\big)$ for $n=0, 1, 2, \ldots$. Consequently if $C_0^* f=\lambda f$, then $f(n) = \lambda(n+1)\big(f(n)-f(n+1)\big)$ or $\lambda(n+1)f(n+1) = \big(\lambda(n+1)-1\big)f(n)$. It follows that 0 is not a proper value of C_0^* (if $\lambda=0$, then $f(n)=0$ for all n), and it follows also that $f(n+1) = \left(1-\dfrac{1}{\lambda(n+1)}\right)f(n)$. This implies that if $n\geqq 1$, then

$$f(n) = \prod_{j=1}^{n}\left(1-\frac{1}{j\lambda}\right)f(0),$$

and we can conclude, even before we know which values of λ can be proper values of C_0^*, that all the proper values are simple.

Suppose now that $|1-\lambda|<1$, or, equivalently, that $\operatorname{Re}\dfrac{1}{\lambda}>\dfrac{1}{2}$. It is convenient to rewrite the condition once more; if $\mu=\dfrac{1}{\lambda}$, then the condition is that $2\operatorname{Re}\mu = 1+\varepsilon$ for some positive number ε. Our task is to prove that if this condition is satisfied, and if

$$f(n) = \prod_{j=1}^{n}\left(1-\frac{\mu}{j}\right),$$

for $n\geqq 1$, then $f\in l^2$. Since

$$\left|1-\frac{\mu}{j}\right|^2 = 1-\frac{2\operatorname{Re}\mu}{j}+\frac{|\mu|^2}{j^2} = 1-\frac{1+\varepsilon}{j}+\frac{|\mu|^2}{j^2} \leqq \exp\left(\frac{|\mu|^2}{j^2}-\frac{1+\varepsilon}{j}\right),$$

it follows that

$$|f(n)|^2 \leqq \frac{\exp\left(|\mu|^2 \sum_{j=1}^{n} \frac{1}{j^2}\right)}{\exp\left((1+\varepsilon) \sum_{j=1}^{n} \frac{1}{j}\right)} < \frac{c}{\exp((1+\varepsilon)\log n)} = \frac{c}{n^{1+\varepsilon}},$$

where $c = \exp\left(|\mu|^2 \sum_{j=1}^{\infty} \frac{1}{j^2}\right)$. This completes the proof of (4). (We note in passing that if f is a proper vector of C_0^* with proper value λ, then $\sum_{n=0}^{\infty} f(n) z^n = (1-z)^{\frac{1}{\lambda}-1}$ whenever $|z|<1$.)

Since $\|1-C_0\| = 1$, the spectrum of $1-C_0$ is included in the closed disc $\{\lambda: |\lambda| \leqq 1\}$, and, consequently, the spectrum of C_0 is included in the closed disc $\{\lambda: |1-\lambda| \leqq 1\}$. The preceding paragraph implies that the spectrum of $1-C_0^*$ includes the open disc $\{\lambda: |\lambda|<1\}$, and hence that the same is true of the spectrum of $1-C_0$. This, in turn, implies that the spectrum of C_0 includes the open disc $\{\lambda: |1-\lambda|<1\}$, and the proof of (6) is complete.

In view of what was just proved, the proof of (5), and hence of the theorem, can be completed by showing that if $|1-\lambda| = 1$, then λ is not a proper value of C^*, or, equivalently, $1-\lambda$ is not a proper value of $1-C_0^*$. This, however, is an immediate consequence of (2): if $\|f\|=1$ and $(1-C_0^*)f = (1-\lambda)f$, then $\|(1-C_0^*)f\| = |1-\lambda|$, and therefore $|1-\lambda|$ cannot be equal to 1.

We conclude our discussion of the discrete Cesàro operator by reporting a fact that may not be important but that is at least an interesting curiosity.

Theorem 3. *The operator C_0 is hyponormal, that is, $C_0^*C_0 - C_0C_0^*$ is positive.*

Proof. The matrix $k_0^*k_0$ is "L-shaped", meaning that it is of the form

$$\begin{pmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \\ \alpha_1 & \alpha_1 & \alpha_2 & \\ \alpha_2 & \alpha_2 & \alpha_2 & \\ & & & \ddots \end{pmatrix},$$

with $\alpha_n = \sum_{j=n}^{\infty} \frac{1}{(j+1)^2}$. Since $k_0k_0^*$ is also L-shaped $\left(\text{with } \alpha_n = \frac{1}{n+1}\right)$, and since the difference of two L-sharped matrices is another one, the problem of proving the hyponormality of C_0 reduces to the problem of deciding when an L-shaped matrix is positive. An infinite matrix is positive if and only if all its finite sections have positive determinants; the problem has reduced to the evaluation of the determinant of

$$\begin{vmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \alpha_1 & \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \alpha_2 & \alpha_2 & \alpha_2 & \dots & \alpha_n \\ \vdots & \vdots & \vdots & & \vdots \\ \alpha_n & \alpha_n & \alpha_n & \dots & \alpha_n \end{vmatrix}.$$

This is easy. Subtract the second column from the first, then subtract the third column from the second, and continue this way through the columns. The resulting matrix has the same determinant as the original one and is triangular; its determinant therefore is the product of its diagonal elements. The diagonal elements are $\alpha_0-\alpha_1, \alpha_1-\alpha_2, \ldots, \alpha_{n-1}-\alpha_n$, and α_n. Conclusion: an finite L-shaped matrix is positive if and only if its determining sequence is positive and decreasing. The proof of the theorem is completed by verifying that the sequence $\left\{\sum_{j=n}^{\infty} \frac{1}{(j+1)^2}-\frac{1}{n+1}\right\}$ has these properties.

The finite continuous Cesàro operator

For C_1 the facts are simpler and the proofs are easier than for C_0; to get at those facts, it is convenient to recall a few simple results about unilateral shifts. An operator U on a Hilbert space H is a unilateral shift of multiplicity 1 if H has an orthonormal basis $\{e_0, e_1, e_2, \ldots\}$ such that $Ue_n=e_{n+1}$, $n=0, 1, 2, \ldots$. A unilateral shift of multiplicity m (here m can be any cardinal number, finite or infinite) is the direct sum of m unilateral shifts of multiplicity 1. Each unilateral shift is an isometry, and so therefore is the direct sum of a unilateral shift and a unitary operator. Conversely, every isometry is a direct sum of a unilateral shift and a unitary operator, it being understood that either summand may be absent. If U is an isometry, then U^*U-UU^* is the projection on the co-range of U (the orthogonal complement of the range of U), and consequently the rank of U^*U-UU^* (the co-rank of U) is the multiplicity of the shift component of U.

If U is a unilateral shift, then the spectrum of U is the closed unit disc, the point spectrum of U is empty, and the point spectrum of U^* is the open unit disc. Each number in the open unit disc is a proper value of U^* of multiplicity equal to the multiplicity of U. The proper vectors of U^* form a total set (that is, they span the entire underlying Hilbert space). All these facts are known; see [1, 2, 4].

There are several ways of characterizing simple unilateral shifts (that is, unilateral shifts of multiplicity 1). For our purposes the most convenient one is this: an operator U is a simple unilateral shift if and only if (1) U is an isometry, (2) the co-rank of U is 1, and (3) U^* has a total set of proper vectors with proper values of modulus strictly less than 1. Indeed, a unilateral shift has these three properties. If, conversely, U is an operator satisfying (1), (2), and (3), then, by (1), it is he direct sum of a unilateral shift and a unitary operator, and, by (2), its shift component is simple. It remains only to use (3) to prove that its unitary component is absent. Suppose therefore that W is a unitary direct summand of U. If $U^*f=\lambda f$ with $|\lambda|<1$, and if g is the component of f in the domain of W, then $W^*g=\lambda g$; since W^* is unitary, it follows that $g=0$. Thus each proper vector of U^* corresponding to a proper value of modulus strictly less than 1 belongs to the domain of the shift component of U; if such vectors span the whole space, then the unitary component of U cannot be present.

Theorem 4. *The operator* $1-C_1^*$ *is a simple unilateral shift.*

Proof. Since C_1 is given by the kernel k_1, where $k_1(x, y)=1/x$ if $0<y\leqq x$ and $k_1(x, y)=0$ otherwise, it follows that C_1^* is given by the kernel k_1^*, where

$$k_1^*(x, y) = \begin{cases} 0 & \text{if } 0<y\leqq x, \\ \dfrac{1}{y} & \text{if } 0<x<y. \end{cases}$$

In other words if $f\in L^2(0, 1)$, then

$$(C_1^* f)(x) = \int_x^1 \frac{1}{y} f(y)\, dy.$$

The operator $C_1 C_1^*$ is given by the kernel

$$\int_0^1 k_1(x, u) k_1^*(u, y)\, du = \int_0^{\min(x,y)} \frac{1}{x}\frac{1}{y}\, du = \frac{\min(x, y)}{xy}.$$

Since

$$k_1(x, y)+k_1^*(x, y) = \begin{cases} \dfrac{1}{x} & \text{if } 0<y\leqq x, \\ \dfrac{1}{y} & \text{if } 0<x<y, \end{cases}$$

it follows that $C_1 C_1^* = C_1 + C_1^*$, and hence that

$$(1-C_1)(1-C_1^*) = 1.$$

Conclusion: $1-C_1^*$ is an isometry.

If we write $1-C_1^* = U$, then $U^*U-UU^* = C_1C_1^*-C_1^*C_1$. Since $C_1^*C_1$ is given by the kernel

$$\int_0^1 k_1^*(x, u) k_1(u, y)\, du = \int_{\max(x,y)}^1 \frac{du}{u^2} = \frac{1}{\max(x, y)}-1,$$

it follows that the kernel of $C_1C_1^*-C_1^*C_1$ is the constant function 1. Conclusion: the co-rank of $1-C$ is equal to 1.

Before completing the proof of the theorem, we remark on the kernel techniques used in the proof so far. Since the kernels in question are neither in L^2 (that is, the operators are not in the Hilbert—Schmidt class), nor symmetric (the two textbook cases), it is not quite automatic that if an operator is given by a kernel, then its adjoint is given by the conjugate transpose kernel, and that the product of two operators given by kernels is given by the product kernel. Since, however, the kernels k in question (that is, k_1 and k_1^*) have positive values, and have the property that if f and g are in L^2, then the function given on the unit square by $k(x, y) f(x) g(y)$ is in L^1, no unboundedness or infinity pathology can occur; the necessary changes in the order of integration are immediate consequences of FUBINI's theorem.

To complete the proof of the theorem it is sufficient to show that $1-C_1$ has a total set of proper vectors corresponding to proper values of modulus strictly less than 1. This is trivial modulo the Weierstrass approximation theorem. If $f_n(x)=x^n$, $n=0, 1, 2, \ldots,$ then the set $\{f_0, f_1, f_2, \ldots\}$ is total in $L^2(0, 1)$. Since $(C_1 f_n)(x) =$ $= \frac{1}{x}\int_0^x y^n\,dy = \frac{x^n}{n+1} = \frac{1}{n+1} f_n(x)$, it follows that $(1-C_1)f_n = \left(1-\frac{1}{n+1}\right)f_n$, and the proof is complete.

It may be worth while to remark that Theorem 4 implies that all the spectral assertions of Theorem 2 ((3), (4), (5), and (6)) remain true, word for word, if in their statement C_0 is replaced by C_1^*. The norm assertion (1) is also invariant under this change; the only part of the theorem that changes is (2). Since $1-C_1^*$ is an isometry, $\|(1-C_1^*)f\| = \|f\|$ always and $\|(1-C_1)f\| = \|f\|$ often. What can be said, however, is that if $\|f\|=1$, then $\|C_1 f\|<2$ and $\|C_1^* f\|<2$. This follows either by an examination of the cases of equality in the Schur test, or by a direct argument valid for isometries with no proper values.

Here is another useful comment about unilateral shifts, and hence about $1-C_1^*$. The basis that a simple unilateral shift shifts is uniquely determined to within a multiplicative constant. The reason is that the co-range is one-dimensional and e_0 is in the co-range. Since the projection on the co-range of $1-C_1^*$ is $C_1C_1^*-C_1^*C_1$, and since, as we have seen, this projection is given by the kernel that is identically 1, it follows that the co-range of $1-C_1^*$ is the set of all constant functions. The most natural choice for e_0 is the constant function 1. Once e_0 is chosen, the other terms of the shifted basis are determined; they are the successive images of e_0 under iterations of $1-C_1^*$.

There is another approach to Theorem 4, more analytic than the one given above; we proceed to sketch it. If $U = 1-C_1^*$ and $f_\alpha(x)=x^\alpha$ whenever $\operatorname{Re}\alpha>-\frac{1}{2}$, then $U^*f_\alpha = \frac{\alpha}{\alpha+1}f_\alpha$. A change of parameters is convenient: if $\beta=\bar\alpha+\frac{1}{2}$ and $g_\beta=f_{\bar\beta-\frac{1}{2}}$ whenever $\operatorname{Re}\beta>0$, then $U^*g_\beta=\overline{\varphi(\beta)}g_\beta$, where $\varphi(\beta) = \frac{\beta-\frac{1}{2}}{\beta+\frac{1}{2}}$.

By means of these proper vectors, the operator U can be represented as a multiplication on a Hilbert space of analytic functions on the right half plane, as follows. For f in $L^2(0, 1)$ define $\hat f$ by

$$\hat f(\beta) = (f, g_\beta) = \int_0^1 f(t)\,t^{\beta-\frac{1}{2}}\,dt;$$

the transform of U by the mapping $f\to\hat f$ is multiplication by φ. Indeed,

$$(Uf)^\wedge(\beta) = (Uf, g_\beta) = (f, U^*g_\beta) = \Phi(\beta)\hat f(\beta).$$

Making the change of variables $t = e^{-u}$ $(0<u<\infty)$, we obtain

$$\hat f(\beta) = \int_0^\infty f(e^{-u})\,e^{-u/2}\,e^{-u\beta}\,du = \int_0^\infty g(u)\,e^{-u\beta}\,du,$$

where g is the element of $L^2(0, \infty)$ defined by $g(u)=f(e^{-u})e^{-u/2}$. Thus the space of functions $\hat{f}$ is the space of Laplace transforms of functions in $L^2(0, \infty)$. By the Paley—Wiener theorem [6, Chapter VIII] this is precisely the space H^2 of the right half plane, and therefore the preceding paragraph exhibits U as multiplication by φ on that H^2 space. Switching to the unit disc via the conformal mapping $w=\varphi(z)$, we obtain a representation of U as multiplication by the independent variable on H^2 of the disc, and Theorem 4 follows.

We conclude our discussion of the finite continuous Cesàro operator by mentioning a curious by-product of Theorem 4. One of our earlier proofs of that theorem made use of the completeness of the set of Laguerre functions in $L^2(0, \infty)$. The proof actually offered above is independent of such considerations; since it turns out that our earlier argument is reversible, Theorem 4 can be used to prove that the Laguerre functions span $L^2(0, \infty)$. Here is how it goes. If $f \in L^2(0, 1)$, write

$$(Tf)(x) = f(e^{-x})e^{-x/2}$$

for $0 < x < \infty$, and verify that T is an isometry from $L^2(0, 1)$ onto $L^2(0, \infty)$. Transform the shift $1 - C_1^*$ by T; that is, consider on $L^2(0, \infty)$ the operator $V = T(1-C_1^*)T^{-1}$. If $f \in L^2(0, \infty)$, then Vf can be calculated explicitly:

$$(Vf)(x) = f(x) - e^{-x/2}\int_0^x f(y)e^{y/2}\,dy.$$

If, as usual, the Laguerre polynomials are defined by

$$L_n(x) = \frac{1}{n!}e^x \frac{d^n}{dx^n}(x^n e^{-x}),$$

and the Laguerre functions by

$$f_n(x) = e^{-x/2}L_n(x), \qquad n=0, 1, 2, \ldots,$$

then the f_n's form an orthonormal set in $L^2(0, \infty)$. A straightforward argument, based on the standard identity

$$L_n(x) = \frac{d}{dx}(L_n(x) - L_{n+1}(x))$$

(see [8, Chapter VI]) implies that $Vf_n = f_{n+1}$. Since $Te_0 = f_0$, it follows that $Te_n = f_n$ for $n=0, 1, 2, \ldots$, and the completeness of the f_n's follows from that of the e_n's.

The infinite continuous Cesàro operator

We shall get at the facts about C_∞ by reducing its study to that of C_1. It is convenient to begin by establishing a simple result about the relation between unilateral shifts and bilateral shifts. An operator W on a Hilbert space K is a simple bilateral shift if K has an orthonormal basis $\{\ldots, e_{-2}, e_{-1}, e_0, e_1, e_2, \ldots\}$ such that $We = e_{n+1}$ for all n. It follows from this definition that a simple bilateral shift is a unitary operator. If H is the span of $\{e_0, e_1, e_2, \ldots\}$, then H is invariant under

W and the restriction of W to H is a unilateral shift. If R is the operator on K such that $Re_n = e_{-n-1}$ for all n, then R is a symmetry (a unitary involution). The symmetry R is related to the shift W in the following three ways:

$$(1)\quad Re_0 = W^{-1}e_0, \qquad (2)\quad RH = H^\perp, \qquad (3)\quad RW = W^{-1}R.$$

What makes these assertions important is that they serve to characterize simple bilateral shifts, in the following sense. Suppose that K is a Hilbert space, W is a unitary operator on K, R is a symmetry on K, H is a subspace of K invariant under W, and e_0 is a vector in H. If the vectors $W^n e_0$, $n = 0, 1, 2, \ldots$, form an orthonormal basis for H, and if the conditions (1), (2), and (3) are satisfied, then W is a simple bilateral shift.

The proof is straightforward. We begin by writing $e_n = W^n e_0$ for all n $(=0, \pm 1, \pm 2, \ldots)$. If n and m are arbitrary integers, find a positive integer j such that both $n+j$ and $m+j$ are positive; it follows that

$$(e_n, e_m) = (W^n e_0, W^m e_0) = (W^{n+j}e_0, W^{m+j}e_0) = (e_{n+j}, e_{m+j}) = \delta_{n+j,m+j} = \delta_{nm},$$

and hence that the e_n's form an orthonormal set in K. By assumption $\{e_0, e_1, e_2, \ldots\}$ spans H; it follows that $\{Re_0, Re_1, Re_2, \ldots\}$ spans $H^\perp$. Since $Re_n = RW^n e_0 = W^{-n}Re_0 = W^{-n}W^{-1}e_0 = e_{-n-1}$, it follows that $\{e_{-1}, e_{-2}, e_{-3}, \ldots\}$ spans $H^\perp$, and hence that the e_n's form an orthonormal basis for K. Since the definition of the e_n's makes it obvious that W shifts them, the proof of the characterization of simple bilateral shifts is complete.

Theorem 5. *The operator $1 - C_\infty^*$ is a simple bilateral shift.*

Proof. We apply the preceding characterization of simple bilateral shifts with $K = L^2(0, \infty)$, $W = 1 - C_\infty^*$, and

$$(Rf)(x) = -\frac{1}{x} f\left(\frac{1}{x}\right)$$

whenever $f \in K$. The role of H is played by those elements of K that vanish on $(1, \infty)$, and the role of e_0 is played by the characteristic function of $(0, 1)$. We observe that H differs from $L^2(0, 1)$ in notation only.

If $f \in K$, then

$$(Wf)(x) = f(x) - \int_x^\infty \frac{1}{y} f(y)\,dy$$

for $0 < x < \infty$. With this explicit representation of W, the verifications needed to justify the application of the characterization theorem for bilateral shifts become a matter of routine integrations. They are not only routine, but they are almost identical with the integrations indicated in our study of C_1. (Note that if H is identified with $L^2(0, 1)$, then the restriction of W to H must be identified with $1 - C_1^*$.) With these remarks we consider the proof of Theorem 5 complete.

It follows from Theorem 5 (just as the corresponding facts for C_1 followed from Theorem 4) that $\|1 - C_\infty\| = 1$ and $\|C_\infty\| = 2$; if $\|f\| = 1$, then $\|C_\infty f\| < 2$ and

$\|C_\infty^* f\| < 2$. Using in addition well known (and easily recaptured) facts about the spectrum of a bilateral shift, we obtain the following description of the spectrum of C_∞: the point spectra of both C_∞ and C_∞^* are empty, and the spectrum of C_∞ is the circle $\{\lambda : |1-\lambda| = 1\}$.

References

[1] A. BEURLING, On two problems concerning linear transformations in Hilbert space, *Acta Math.*, **81** (1949), 239—255.
[2] A. BROWN, On a class of operators, *Proc. Amer. Math. Soc.*, **4** (1953), 723—728.
[3] L. DE BRANGES, Some Hilbert spaces of entire functions. III, *Transactions Amer. Math. Soc.*, **100** (1961), 73—115.
[4] P. R. HALMOS, Shifts on Hilbert spaces, *J. reine angew. Math.*, **208** (1961), 102—112.
[5] G. H. HARDY, J. E. LITTLEWOOD, and G. PÓLYA, *Inequalities* (Cambridge, 1934).
[6] K. HOFFMAN, *Banach spaces of analytic functions* (Englewood Cliffs, 1962).
[7] I. SCHUR, Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen, *J. reine angew. Math.*, **140** (1911), 1—28.
[8] F. G. TRICOMI, *Vorlesungen über Orthogonalreihen* (Berlin, 1955).

UNIVERSITY OF MICHIGAN

(Received September 14, 1964)

An embedding theorem for some countable groups

By L. G. KOVÁCS and B. H. NEUMANN in Canberra (Australia)

Every countable soluble group can be embedded in a soluble 2-generator group, the solubility length increasing by no more than 2 in the process: this was shown in [5]. We here extend this result to some of the transfinite generalizations of soluble groups. The method of [5] has to be modified to do this, firstly as in [4] and secondly as in HALL's paper [1].

We use the following notation and terminology. An ascending series of subgroups of a group G is a family $\{L_\lambda\}_{0 \leqq \lambda \leqq \sigma}$ of subgroups of G indexed by the set of ordinals less than or equal to the ordinal σ, and such that $L_0 = \{1\}$ and, for $0 < \lambda \leqq \sigma$

(1) $$L_\lambda = \bigcup_{\mu < \lambda} L_{\mu+1}.$$

[This condition ensures that $L_\mu \leqq L_\lambda$ whenever $\mu \leqq \lambda$, and simultaneously that L_λ is the union of its predecessors when λ is a limit ordinal.] If each L_λ is normal in its successor $L_{\lambda+1}$, or even in G, the series is called "normal" or "invariant", respectively. If for $0 \leqq \lambda < \sigma$

$$[L_{\lambda+1}, L_{\lambda+1}] \leqq L_\lambda, \quad \text{or even} \quad [G, L_{\lambda+1}] \leqq L_\lambda,$$

where $[A, B]$ stands for the mutual commutator group of A and B, then the series is called "soluble" or "central", respectively. A soluble series is necessarily normal, and a central series invariant.

If G has a soluble series with $L_\sigma = G$, then G is defined to be an SN^*-group; if the soluble series can be chosen invariant, then G is an SI^*-group; if G has a central series with $L_\sigma = G$, then G is a ZA-group. The ordinal σ is called a "length" of G — we do not assume it chosen minimal, and if G has SN^*-length or SI^*-length or ZA-length σ, then it has also every greater length.

We shall prove the following theorem.

Theorem. *Every countable SI^*-group G of length σ can be embedded in a 2-generator SI^*-group of length $\sigma+2$.*

The method of proof yields rather more than the theorem. To every countable group G, we contruct a 2-generator group H which embeds it. The new feature of H is that its second derived group is contained in a certain interdirect power N_σ of G. Let $\mathfrak{C}$ be a class of groups which is closed under the operations of taking subgroups and taking interdirect powers like N_σ. (The reader has to refer to the first paragraph of the proof: an interdirect power F is selected there, and N_σ is a restricted direct power of F.) It follows from our construction that *every countable group in $\mathfrak{C}$ can be embedded in a 2-generator group whose second derived group is in $\mathfrak{C}$*. Some examples of classes which satisfy the conditions on $\mathfrak{C}$ are those of SN^*-groups, ZA-groups, locally nilpotent groups, locally finite groups, periodic groups, etc. In particular, it follows that *every countable SN^*-group of length σ is embeddable in a 2-generator SN^*-group of length $\sigma+2$.*

We mention an easy consequence of the theorem itself:

Corollary. *There exist SI^*-groups that are not locally soluble.*

This fact was pointed out by HALL in [2]; in the present context it follows by applying the theorem to a countable insoluble SI^*-group G, for instance to one of the characteristically simple groups of MCLAIN [3].

Proof of the Theorem. In addition to the notation introduced above, we also use the definitions and notation of [5]. In the complete wreath product $P=G \text{ Wr } C$ of the given SI^*-group G and an infinite cyclic group C generated by an element c, we single out a subgroup that contains the restricted wreath product $G \text{ wr } C$, but is not much larger. In the base group of P, that is the cartesian power G^C consisting of all functions on C to G, we single out those functions f that are constant for all sufficiently large positive powers of c, and also for all sufficiently large negative powers of c, the constant in this latter case being 1; thus we consider those f to which there is an integer $p \geqq 0$, depending on f, such that

$$f(c^n)=1 \text{ when } n<-p, \quad f(c^n)=f(c^{n+1}) \text{ when } n>p.$$

These functions form a subgroup F of G^C, and F is normalized by C. We put $FC=P^0$.

The cartesian powers L_λ^C are normal subgroups of G^C, but they will not in general form an ascending series in G^C, as the analogue of (1) may fail for limit ordinals λ. However, if we put, for $0 \leqq \lambda \leqq \sigma$,

$$M_\lambda = F \cap L_\lambda^C,$$

so that M_λ consists of those functions $f \in F$ that take values in L_λ, then each M_λ is a normal subgroup of $M_\sigma = F$ and indeed of P^0, and in fact $\{M_\lambda\}_{0 \leqq \lambda \leqq \sigma}$ is an ascending soluble invariant series of P^0. We omit the easy verification. If we put $M_{\sigma+1}=P^0$, then the thus augmented series shows that P^0 is an SI^*-group of length $\sigma+1$.

Next we take an infinite cyclic group B generated by an element b and form the complete wreath product

$$Q = P^0 \text{ Wr } B.$$

This contains in its base group P^{0B} the direct powers N_λ of the M_λ, that is the functions on B to M_λ with finite support. These are easily seen to form an ascending soluble invariant series $\{N_\lambda\}_{0 \leqq \lambda \leqq \sigma+1}$ in Q.

We now use the assumption that G is countable, and generate it by a family $\{g_i\}_{i \in I}$ of elements indexed by the set I of positive integers. To these we define elements $k_i \in F$ by

$$k_i(c^n) = 1 \text{ when } n<0, \quad k_i(c^n) = g_i^{-1} \text{ when } n \geqq 0.$$

Put $g_{i1} = [k_i, c]$; then

$$g_{i1}(1) = g_i, \quad g_i(c^n) = 1 \text{ when } n \neq 0.$$

Thus the family $\{g_{i1}\}_{i \in I}$ generates the coordinate subgroup G_1 of G^C; clearly $G_1 \cong G$. Next we define an element $a \in P^{0B}$ by

$$a(b) = c, \quad a(b^{2^i}) = k_i \text{ when } i \in I,$$

$$a(b^n) = 1 \text{ when } n \text{ is not a power of } 2.$$

Let H be the subgroup of Q generated by a and b, and let A be the normal closure of a in H. Then A is generated by the conjugates

$$a^{b^n} = a_n,$$

say, of a, where n ranges over all integers.

We now show that the derived group A' of A is contained in N_σ. First we remark that A' is generated by all commutators $[a_m, a_0]$ and their conjugates under powers of b; and as b normalizes N_σ, it suffices to show that every $[a_m, a_0]$ lies in N_σ. Now $[a_m, a_0]$ is a function on B to P^0, and we compute its value at b^n:

$$[a_m, a_0](b^n) = [a_m(b^n), a_0(b^n)] = [a(b^{n-m}), a(b^n)];$$

this is 1 unless $n-m$ and n are distinct powers of 2, say $n-m = 2^i$, $n = 2^j$, with i, j non-negative integers. In this case $m = 2^j - 2^i$, and to any one m there is at most one such pair i, j. Thus the support of $[a_m, a_0]$ consists of at most one element of B; it only remains to show that the one non-trivial value of $[a_m, a_0]$, if it has one at all, lies in $M_\sigma = F$. Now if $m = 2^j - 2^i \neq 0$, then

$$\begin{aligned}[a_m, a_0](b^{2^j}) = [a(b^{2^i}), a(b^{2^j})] &= g_{j1}^{-1} \quad \text{if} \quad i = 0,\\ &= g_{i1} \quad \text{if} \quad j = 0,\\ &= [k_i, k_j] \quad \text{if} \quad i \neq 0,\ j \neq 0.\end{aligned}$$

These values all lie in F, and it follows that $A' \leqq N_\sigma$ as claimed.

Incidentally the above argument also shows how G can be embedded in H; for if we put, for $i \in I$,

$$h_i = [a_{1-2^i}, a_0],$$

then

$$h_i(b) = g_{i1}, \quad h_i(b^n) = 1 \text{ when } n \neq 1;$$

hence the subgroup of H generated by $\{h_i\}_{i \in I}$ is isomorphic to G_1 and thus to G.

Finally we put $K_\lambda = H \cap N_\lambda$ for $0 \leqq \lambda \leqq \sigma$. Then, as $\{N_\lambda\}_{0 \leqq \lambda \leqq \sigma}$ is an ascending soluble invariant series of Q, also $\{K_\lambda\}_{0 \leqq \lambda \leqq \sigma}$ is an ascending soluble invariant series of H. Adding $K_{\sigma+1} = A$ and $K_{\sigma+2} = H$ to this series, we obtain an ascending soluble invariant series that terminates with H itself; for as we have just seen, $A' \leqq N_\sigma$ and thus also $K'_{\sigma+1} \leqq K_\sigma$; and obviously also $H' \leqq A$. It follows that H is an SI^*-group of length $\sigma + 2$, and the Theorem is proved.

References

[1] P. Hall, The Frattini subgroups of finitely generated groups, *Proc. London Math. Soc.*, (3) **11** (1961), 327–352.
[2] P. Hall, On non-strictly simple groups, *Proc. Cambridge Philos. Soc.*, **59** (1963), 531–553.
[3] D. H. McLain, A characteristically simple group, *Proc. Cambridge Philos. Soc.*, **50** (1954), 641–642.
[4] B. H. Neumann, Embedding theorems for ordered groups, *J. London Math. Soc.*, **35** (1960), 503–512.
[5] B. H. Neumann and Hanna Neumann, Embedding theorems for groups, *J. London Math. Soc.*, **34** (1959), 465–479.

THE AUSTRALIAN NATIONAL UNIVERSITY

(Received April 10, 1964)

On the existence of Baur-soluble groups of arbitrary height

By L. G. KOVÁCS and B. H. NEUMANN in Canberra (Australia)

A hierarchy of generalized soluble groups has been defined by HERIBERT BAUR in [1]: abelian groups are to be termed Baur-soluble of height 0; for an ordinal h other than 0, a group is to be called Baur-soluble of height h if it is not Baur-soluble of height less than h and has an invariant system (in the sense of KUROSH [3]) whose factors are all Baur-souble of height less than h. Thus Baur-soluble groups of height 1 are precisely the non-abelian SI-groups. It is easy to show, by induction on h, that

(1) subgroups and factor groups of Baur-soluble groups of height h are Baur-soluble of height at most h.

A further simple observation is that

(2) every restricted direct product of Baur-soluble groups is Baur-soluble, and its height is the least upper bound of the heights of the factors.

The purpose of this note is to demontstrate the existence of Baur-soluble groups of height h, for every ordinal h. In fact, the following result will be proved, by induction on h:

Theorem. *There exist perfect Baur-soluble groups of height h, for every ordinal h.* (A group is called perfect if it is its own commutator subgroup.)

Proof. The initial step: for $h=0$, the unit group is the only example. Observation (2) and the fact that restricted direct products of perfect groups are perfect take care of the limit step in the induction.

It remains *to construct, from any given perfect Baur-soluble group G of height h, a perfect Baur-suloble group of height* $h+1$. The construction is based on wreath powers, which were introduced and studied by HALL in [2]; his notation is adopted throughout. Important use is made of a torsion free, characteristically simple group A constructed by MCLAIN in [5]. This group has the additional property that it is the product of its abelian normal subgroups. Thus, firstly, A is a perfect SI-group, and secondly, A is locally nilpotent. Therefore, firstly, A is a perfect Baur-soluble group of height 1, and so there is no need to consider the case $h=0$. Secondly, as a torsion free locally nilpotent group, A can be (fully) ordered (cf. MAL'CEV [4], or Corollary 6. 2 of B. H. NEUMANN [6]). Let Λ be the ordered set of elements of A, and let W be the wreath power Wr H^{Λ}. For each element λ of Λ, HALL defines (in section 2. 4 of [2]) a subgroup D_λ of W and shows that

(3) each D_λ is a direct product of isomorphic copies of H;

(4) the D_λ generate W; and

(5) D_λ normalizes D_μ whenever $\mu \in \Lambda$, $\mu \leqq \lambda$. Define D_Γ, for each lower section Γ of Λ, to be the subgroup of W generated by the D_λ with $\lambda \in \Gamma$. Contrary to HALL's convention, allow the empty set $\varnothing$ as a lower section of Λ, and put $D_\varnothing$ equal the unit subgroup. It follows from (4) and (5) that the D_Γ are all normal subgroups of W. As unions of ascending chains of lower sections, as well as intersections of any number of lower sections, are again lower sections, it follows that the D_Γ form an invariant system of W. Let D_Δ/D_Γ be a factor of this system; then no member of the system can lie properly between D_Γ and D_Δ, and so no lower section properly contained in Δ can properly contain Γ. From this it is easy to deduce that Δ contains just one element λ which is not contained in Γ. Hence $D_\Delta = D_\Gamma D_\lambda$; so that D_Δ/D_Γ is isomorphic to a factor group of D_λ. Now (1), (2) and (3) give that each factor of this invariant system of W is Baur-soluble of height at most h, and therefore W is Baur-soluble of height at most $h+1$. On the other hand, as W is generated by the perfect Baur-soluble groups D_λ of height h, W is perfect, and its height is at least h.

If the height of W is $h+1$, the construction is complete. Assume that W is of height h. Let the elements of A act on Λ as right translations; then A becomes a transitive and hence irreducible group of order-preserving permutations of Λ. Take G to be the natural split extension of W by A. (In [2], HALL takes A to be the group of *all* order-preserving permutations of Λ, but the omission of "all" does not effect his subsequent results.) Since $G/W \cong A$ and A is Baur-soluble of height 1, G is Baur-soluble of height at most $h+1$; also, as W and A are perfect, so is G. It follows from Theorem D of HALL [2] that W is a minimal normal subgroup of G; therefore every invariant system of G has a factor which contains an isomorphic copy of W and therefore cannot be Baur-soluble of height less than H. Consequently, the height of G cannot be less than $h+1$. This completes the construction in case W is of height h.

References

[1] HERIBERT BAUR, Ein Kommutativitätskriterium für unendliche auflösbare Gruppen, *Arch. Math.*, **11** (1960), 176–182.

[2] P. HALL, Wreath powers and characteristically simple groups, *Proc. Cambridge Philos. Soc.*, **58** (1962), 170–184.

[3] A. G. KUROSH, *Theory of groups*, vol. 2 (New York, 1956).

[4] A. I. MAL'CEV, On the completion of group order, *Trudy Mat. Inst. Steklov*, **38** (1951), 173–175 (Russian).

[5] D. H. MCLAIN, A characteristically-simple group, *Proc. Cambridge Philos. Soc.*, **50** (1954), 641–642.

[6] B. H. NEUMANN, An embedding theorem for algebraic systems, *Proc. London Math. Soc.*, (3) **4** (1954), 138–153.

THE AUSTRALIAN NATIONAL UNIVERSITY

(Received April 10, 1964)

On products of normal subgroups

By GILBERT BAUMSLAG, L. G. KOVÁCS and B. H. NEUMANN in Canberra (Australia) *)

The purpose of this note is to demonstrate that the product of two "good" normal subgroups of a group need not be "good": we do this for the cases when "good" is interpreted, in turn, as locally soluble, locally SI^*, $\overline{SI}$, locally SN^*, locally residually nilpotent, and residually nilpotent. The first four interpretations answer questions recorded in section 2. 4 of PLOTKIN's survey [3][1]); the fifth answers question 13. 1. 3 of [3]; and the last contradicts an assertion of SESEKIN ([4], part of Corollary 1 to Lemma 8) quoted in [3] (in the paragraph immediately preceding 13. 1. 3). In fact, we construct two examples which show even more:

Theorem 1. *There exists a finitely generated group which is the product of two locally soluble normal subgroups but is neither an SI-group nor a radical group* (in the sense of PLOTKIN [3]; note that, according to § 15 of [3], it follows that the group is not an SN^*-group).

Theorem 2. *There exists a finitely generated* (*torsion-free, metabelian*) *group which is the product of two residually nilpotent normal subgroups but is not residually nilpotent.*

The proofs depend on the following:

Lemma. *To each group H which is a split extension of a group G by an abelian group A, it is possible to construct a group H^* which contains a subgroup isomorphic to H and is the product of two normal subgroups isomorphic to the restricted* (*standard*) *wreath product G* wr *A. Moreover, if H is finitely generated, then so is the corresponding H^*.*

A similar construction has been used independently by HALL (unpublished) for dealing with some of PLOTKIN's questions which are answered by Theorem 1. A result similar to Theorem 2 can also be deduced from a theorem of HALL and HARTLEY (to appear) according to which every group is embeddable in a suitable product of two normal free subgroups.

Proof of the Lemma. Let H be a split extension of a group G by an abelian group A; it can be assumed that G is a normal subgroup of H, $A \cap H = 1$, and there

*) The first author acknowledges support from the National Science Foundation of the U.S.A.

1) We note that the remaining question of this type in 2. 4 of [3], namely the question relating to the class $\overline{SN}$, has a positive answer: it is straightforward to see that even all extensions of $\overline{SN}$-groups by $\overline{SN}$-groups are $\overline{SN}$-groups.

is a monomorphism $\beta: A \to H$ for which $G \cap A\beta = 1$ and $G(A\beta) = H$. Consider the unrestricted (standard) wreath product P of H and A, and write its base group H^A as the group of functions from A to H with valuewise multiplication. Call K that subgroup of H^A which consists of those functions whose values are all in G and are in fact equal to 1 at all but finitely many elements of A. This K is normal in P, and $KA \cong G \operatorname{wr} A$. For each element a in A, let $a\delta$ be the constant function on A with value $a\beta$; then $\delta: A \to H^A$ is a monomorphism; moreover, $A\delta$ and A generate an abelian subgroup in P. We need next the element f of H^A defined by

$$f(a) = a\beta \qquad \text{for every } a \text{ in } A.$$

Straightforward calculation shows that

$$(*) \qquad f^{-1}af = a\delta \cdot a \qquad \text{for every } a \text{ in } A,$$

so that $A^f \leqq (A\delta)A$, and therefore A and A^f commute elementwise. Consequently, KA and KA^f normalize each other, and so their product H^* is a subgroup of P. As $KA^f = (KA)^f$, we have that H^* is the product of two normal subgroups isomorphic to $G \operatorname{wr} A$. To each element g of G, let the element $g\gamma$ of H^A be defined by

$$(g\gamma)(1) = g, \quad (g\gamma)(a) = 1 \qquad \text{whenever } 1 \neq a \in A.$$

Then $\gamma: G \to H^A$ is a monomorphism; moreover, $G\gamma \leqq K$. Each element of H is uniquely a product $g(a\beta)$ with $g \in G$, $a \in A$; the mapping $\alpha: g(a\beta) \to (g\gamma)(a\delta)$ is therefore well defined; in fact, it is a monomorphism of H into H^A. As $G\gamma \leqq K$, and as (*) shows that $A\delta \leqq AA^f$, it follows that $H\alpha \leqq H^*$: so H has a subgroup isomorphic to H. Finally, suppose that H is finitely generated, and note that in this case so is A. It is easily seen that $G\gamma$ and A generate KA; hence H^* is generated by G_γ, A, and A^f. By (*), A and A^f generate the same subgroup as A and $A\delta$; hence H^* is generated by $G\gamma$, $A\delta$, and A. The subgroup generated by $G\gamma$ and $A\delta$ is precisely $H\alpha$. Thus we conclude that H^* is generated by the finitely generated subgroups $H\alpha$ and A, and so H^* itself is finitely generated.

Proof of Theorem 1. We use the terminology and results of HALL [1]. Let G be the wreath power $\operatorname{Wr} C^Z$ of an infinite cyclic group C corresponding to the naturally ordered set Z of rational integers, and let A be the group of all order-preserving permutations of Z. Then A is again an infinite cyclic group, and there is a natural split extension H of G by A. Clearly, this H is finitely generated. According to Theorem D of HALL [1], G' is a minimal normal subgroup of H. It is easy to see that G' is not locally nilpotent; consequently, no group containing H can be an SI-group or a radical group. On the other hand, G is locally soluble and therefore so is $G \operatorname{wr} A$. The corresponding group H^* of the Lemma provides the example required for the theorem.

Proof of Theorem 2. Let G be the group defined on the generators $g_1, g_2, \dots$ by the relations $g_i = g_{i+1}^2$, $i = 1, 2, \dots$; then G is an abelian group of rank 1. Let A be the group of automorphisms of G generated by the automorphism α: $g \to g^2$, and let H be the natural split extension of G by A, with a an element from that coset of G in H which corresponds to α. Then g_1 and a generate H; moreover, the relation

$$[g, a] = g \qquad \text{for every } g \text{ in } G$$

shows that G lies in every term of the lower central series of H: consequently, H is not residually nilpotent, and so no group containing H can be residually nilpotent. On the other hand, according to Lemma 14 of HALL [2], $G \operatorname{wr} A$ is residually nilpotent. Thus the corresponding group H^* of the Lemma provides the required example.

References

[1] P. HALL, Wreath powers and characteristically simple groups, *Proc. Cambridge Philos. Soc.*, **58** (1962), 170—184.
[2] P. HALL, On non-strictly simple groups, *Proc. Cambridge Philos. Soc.*, **59** (1963), 531—553.
[3] B. I. PLOTKIN, Generalized soluble and generalized nilpotent groups, *Uspehi Mat. Nauk* (N. S.), **13** (1958), 89—172 [Russian]; *Amer. Math. Soc. Translations*, (2) **17** (1961), 29—117.
[4] N. F. SESEKIN, On locally nilpotent torsion-free groups, *Mat. Sbornik* (N. S.), **32** (74) (1953), 407—442 [Russian].

THE AUSTRALIAN NATIONAL UNIVERSITY

(Received August 12, 1964)

Über die zulässigen Ideale in Szépschen Ringerweiterungen

Von L. C. A. VAN LEEUWEN in Delft (Holland)

Einleitung

In seiner Arbeit [4] hat J. SZÉP den Begriff der *allgemeinen Zerlegung* eines Ringes R eingeführt, d. h. R ist ein Ring mit zwei Unterringen A und B so beschaffen, daß die Beziehungen $R^+ = A^+ \oplus B^+$, $A \cap B = 0$ gelten, wobei durch R^+, A^+ und B^+ der Modul des Ringes R, A und B bezeichnet wird und $\oplus$ das Zeichen der direkten Summe (des Moduls) ist. Ein solcher Ring R wird bezeichnet mit $R = A \dotplus B$. Umgekehrt sind zu beliebigen Ringen A, B diejenige Ringe R zu bestimmen, für die $R = A' \dotplus B'$, $A' \cap B' = 0$, $A' \cong A$, $B' \cong B$ gelten, wobei $\cong$ die Isomorphie bezeichnet. Der Ring R entsteht also aus den Ringen A und B durch *allgemeine Zusammensetzung* und wir werden R eine *Szépsche Erweiterung* von A und B nennen.

Um die Struktur der Szépschen Erweiterungen kennen zu lernen, werden wir in dieser Arbeit eine Klasse von Idealen in solchen Ringen untersuchen. Ist R eine Szépsche Erweiterung von A und B, so entstehen in A bzw. B die von J. SZENDREI eingeführte *zulässige* Ideale in A bzw. B [3]. Es zeigt sich, daß es spezielle zulässige Ideale $T(\mathfrak{a})$ bzw. $T'(\mathfrak{b})$ in B bzw. A gibt, mit denen man eine Klasse von Idealen $(\mathfrak{a}, T(\mathfrak{a}))$ bzw. $(T'(\mathfrak{b}), \mathfrak{b})$ in R bestimmen kann. Dabei sind $\mathfrak{a}$ bzw. $\mathfrak{b}$ zulässige Ideale in A bzw. B. Wir werden zeigen, daß man mit Hilfe der zulässigen Idealen $T(\mathfrak{a})$ bzw. $T'(\mathfrak{b})$ in B bzw. A ein Kriterium dafür gewinnen kann, daß die Zusammensetzung $\mathfrak{a} \dotplus \mathfrak{b}$ in R ein Ideal in R ist. Dabei sind $\mathfrak{a}$ bzw. $\mathfrak{b}$ beliebige zulässige Ideale in A bzw. B (Satz 2 und 2a). Wenn A noch B ein Ideal in $R = A \dotplus B$ ist und A noch B zulässige Unterringe enthalten, so existieren in R nur *echte* Ideale I, für die I^+ subdirekte Summe von A^+ und B^+ ist (Satz 3). Unter gewissen Voraussetzungen bezüglich der Idealen $T(\mathfrak{a})$ bzw. $T'(\mathfrak{b})$ in B bzw. A ist es möglich einen Verbandsisomorphismus zwischen den Verband der zulässigen Idealen in A und denselben Verband in B anzugeben. Auch besteht dann ein Isomorphismus des Verbandes der zulässigen Idealen in A (oder in B) auf den Verband der Ideale der Form $(\mathfrak{a}, T(\mathfrak{a}))$ in R, wo $\mathfrak{a}$ zulässig in A ist (Sätze 4 und 5). Endlich werden wir den Restklassenring $R/(\mathfrak{a}, \mathfrak{b})$ untersuchen, wo $\mathfrak{a}$ bzw. $\mathfrak{b}$ zulässige Ideale in A bzw. B sind und $(\mathfrak{a}, \mathfrak{b})$ ein Ideal in R ist. Man kann zeigen, daß $R/(\mathfrak{a}, \mathfrak{b}) \cong A/\mathfrak{a} \dotplus B/\mathfrak{b}$ (Satz 6), was eine Verallgemeinerung meines Satzes 1 in [1] ist. Umgekehrt, wenn $A/\mathfrak{a}$ bzw. $B/\mathfrak{b}$ zerfallende Everettsche Ringerweiterungen von $\mathfrak{a}$, bzw. $\mathfrak{b}$ sind [2], dann gibt es eine Szépsche Erweiterung R von A und B, in der $\mathfrak{a}^+ \oplus \mathfrak{b}^+$ Ideal ist derart, daß $R/(\mathfrak{a}, \mathfrak{b}) \cong$ $\cong A/\mathfrak{a} \dotplus B/\mathfrak{b}$ (Satz 7).

1. Es seien A und B zwei gegebene (assoziative) Ringe. Das Szépsche Erweiterungsproblem besteht darin, aus den Ringen A und B alle Ringe R mit

$$R^+ = \mathfrak{S}_1^+ \oplus \mathfrak{S}_2^+, \quad \mathfrak{S}_1 \cong A, \quad \mathfrak{S}_2 \cong B, \quad \mathfrak{S}_1 \cap \mathfrak{S}_2 = 0$$

zu bestimmen, wobei $\mathfrak{S}_1, \mathfrak{S}_2$ Unterringe von R sind. Hierbei bezeichnet $\oplus$ die direkte Summe von Moduln, durch R^+ wird der Modul des Ringes R bezeichnet und $\cong$ ist das Zeichen des Isomorphismus.

Die Lösung dieses Problems gewinnt man folgenderweise: In der Menge $R = A \cdot B$ der (geordneten) Paare (a, b) mit $a \in A$ und $b \in B$ definieren wir die Gleichheit, die Addition und die Multiplikation durch die folgenden Relationen:

$$(a, b) = (a', b') \Leftrightarrow a = a', \quad b = b',$$

$$(a, b) + (a', b') = (a + a', b + b'), \tag{A}$$

$$(a, b)(a', b') = (aa' + b_l a' + ab'_r,\ a_l b' + ba'_r + bb'), \tag{B}$$

wobei die Funktionen

$$b_l a,\ ab_r (\in A), \quad a_l b,\ ba_r (\in B) \tag{C}$$

den „Anfangsbedingungen”

$$b_l 0 = 0\, b_r = 0_l\, a = a\, 0_r = 0, \qquad a_l 0 = 0\, a_r = 0_l b = b\, 0_r = 0, \tag{D}$$

unterworfen sind. Die Struktur R ist dann und nur dann ein Ring, wenn

$$\begin{aligned} (b + b')_l a &= b_l a + b'_l a, & a(b + b')_r &= ab_r + ab'_r, \\ (a + a')_l b &= a_l b + a'_l b, & b(a + a')_r &= ba_r + ba'_r, \end{aligned} \tag{1}$$

$$\begin{aligned} (bb')_l a &= b_l(b'_l a), & a(bb')_r &= (ab_r)b'_r, \\ (aa')_l b &= a_l(a'_l b), & b(aa')_r &= (ba_r)a'_r, \end{aligned} \tag{2}$$

$$\begin{aligned} b_l(a + a') &= b_l a + b_l a', & (a + a')b_r &= ab_r + a'b_r, \\ a_l(b + b') &= a_l b + a_l b', & (b + b')a_r &= ba_r + b'a_r, \end{aligned} \tag{3}$$

$$\begin{aligned} b_l(aa') &= (b_l a)a' + (ba_r)_l a', & (aa')b_r &= a(a'b_r) + a(a'_l b)_r, \\ a_l(bb') &= (a_l b)b' + (ab_r)_l b', & (bb')a_r &= b(b'a_r) + b(b'_l a)_r, \end{aligned} \tag{4}$$

$$\begin{aligned} a(b_l a') + a(ba'_r)_r &= (ab_r)a' + (a_l b)_l a', \\ b(a_l b') + b(ab'_r)_r &= (ba_r)b' + (b_l a)_l b', \end{aligned} \tag{5}$$

$$\begin{aligned} (b_l a)b'_r &= b_l(ab'_r), \\ (a_l b)a'_r &= a_l(ba'_r), \end{aligned} \tag{6}$$

für alle $a, a' \in A$ und $b, b' \in B$ erfüllt sind. Diese Ringe sind bis auf Isomorphie die sämtlichen Szépschen Erweiterungen R von A und B. ([3], [4], [5]). Die Elemente $(a, 0)$, $a \in A$ bilden einen Ring $(A, 0) \cong A$, und ebenso ist $(0, B) \cong B$. $(A, 0)$ und $(0, B)$ sind Unterringe von R und $R^+ = (A, 0)^+ \oplus (0, B)^+$.

2. Definition 1. Ein Unterring T von B wird *zulässig* genannt, wenn $a_l b, ba_r \in T$ für alle $b \in T$, $a \in A$ erfüllt ist. Ein Unterring S von A wird zulässig genannt, wenn $b_l a$, $ab_r \in S$ für alle $a \in S$, $b \in B$ erfüllt ist.

Es sei R eine Szépsche Erweiterung von A und B, bezeichnet mit $R = A \dotplus B$ und $\mathfrak{a}$ ein gegebenes zulässiges Ideal in A. (Das Wort „Ideal" hat immer die Bedeutung: „zweiseitiges Ideal"). Wir betrachten die Menge $T(\mathfrak{a})$ von Elementen $b \in B$, für die

$$(7) \qquad b_l a \in \mathfrak{a}, \quad ab_r \in \mathfrak{a}, \quad a(ba'_r)_r \in \mathfrak{a}$$

für alle $a, a' \in A$ erfüllt ist.

Wir beweisen, daß $T(\mathfrak{a})$ *ein Ideal in B ist*. Sind $b_1, b_2 \in T(\mathfrak{a})$, so folgt: $(b_1 - b_2)_l a = b_{1l}a - b_{2l}a \in \mathfrak{a}$ für alle $a \in A$; $a(b_1 - b_2)_r = ab_{1_r} - ab_{2_r} \in \mathfrak{a}$ für alle $a \in A$; $a((b_1 - b_2)a'_r)_r = a(b_1 a'_r - b_2 a'_r)_r = a(b_1 a'_r)_r - a(b_2 a'_r)_r \in \mathfrak{a}$ für alle $a, a' \in A$. Also $b_1 - b_2 \in T(\mathfrak{a})$. Ist $c \in B$, dann ist $cb \in T(\mathfrak{a})$ für jedes Element $b \in T(\mathfrak{a})$. Denn $(cb)_l a = c_l(b_l a) \in \mathfrak{a}$ für alle $a \in A$, weil $\mathfrak{a}$ zulässig ist in A; $a(cb)_r = (ac_r)b_r \in \mathfrak{a}$ für alle $a \in A$, weil $ac_r \in A$ für alle $a \in A$; und $a((cb)a'_r)_r = a(c(ba'_r) + c(b_l a')_r)_r = (ac_r)(ba'_r)_r + a(c(b_l a')_r)_r$ wobei $(ac_r)(ba'_r)_r \in \mathfrak{a}$ für alle $a, a' \in A$ nach (7) und $a(c(b_l a')_r)_r = -a(c_l(b_l a')) + (ac_r)(b_l a') + (a_l c)_l(b_l a') \in \mathfrak{a}$ für alle $a, a' \in A$ da jeder der Summanden zu $\mathfrak{a}$ gehört, also $a((cb)a'_r)_r \in \mathfrak{a}$ für alle $a, a' \in A$. Hieraus folgt $cb \in T(\mathfrak{a})$ für jedes Element $b \in T(\mathfrak{a})$. Ganz ähnlich sieht man ein, daß $bc \in T(\mathfrak{a})$ für jedes Element $b \in T(\mathfrak{a})$. Da c ein beliebiges Element aus B ist, ist damit gezeigt, daß $T(\mathfrak{a})$ ein Ideal in B ist.

$T(\mathfrak{a})$ *ist auch ein zulässiges Ideal in B* nach Definition 1. Dies bedeutet, daß $a_l b, ba_r \in T(\mathfrak{a})$ für alle $b \in T(\mathfrak{a})$, $a \in A$. Nach (7) haben wir zu zeigen, daß $(a_l b)_l a' \in \mathfrak{a}$, $a'(a_l b)_r \in \mathfrak{a}$, $a'((a_l b)a''_r)_r \in \mathfrak{a}$ für alle $a', a'' \in A$ erfüllt ist, und ähnlich für das Element ba_r. Für das Element $a_l b$ gilt: $(a_l b)_l a' = a(b_l a') + a(ba'_r)_r - (ab_r)a' \in \mathfrak{a}$, weil $b_l a' \in \mathfrak{a}$, $a(ba'_r)_r \in \mathfrak{a}$ und $ab_r \in \mathfrak{a}$; $a'(a_l b)_r = (a'a)b_r - a'(ab_r) \in \mathfrak{a}$, weil $a'a \in A$ und $ab_r \in \mathfrak{a}$; $a'((a_l b)a''_r)_r = (a'(a_l b)_r)a'' + (a'_l(a_l b))_l a'' - a'((a_l b)_l a'') \in \mathfrak{a}$, da $a'(a_l b)_r \in \mathfrak{a}$ und $(a_l b)_l a'' \in \mathfrak{a}$ (soeben bewiesen); daraus folgt $a_l b \in T(\mathfrak{a})$. Für das Element ba_r hat man: $(ba_r)_l a' = b_l(aa') - (b_l a)a' \in \mathfrak{a}$; $a'(ba_r)_r = (a'b_r)a + (a'_l b)_l a - a'(b_l a) \in \mathfrak{a}$, weil $a'_l b \in T(\mathfrak{a})$; $a'((ba_r)a''_r)_r = a'(b(aa'')_r)_r \in \mathfrak{a}$; daraus folgt $ba_r \in T(\mathfrak{a})$. Damit ist bewiesen, daß $T(\mathfrak{a})$ ein zulässiges Ideal in B ist.

Ebenso läßt sich zeigen, daß die Menge $T'(\mathfrak{b})$ von Elementen $a \in A$ für die

$$(8) \qquad a_l b \in \mathfrak{b}, \quad ba_r \in \mathfrak{b}, \quad b(ab'_r)_r \in \mathfrak{b}$$

für alle $b, b' \in B$ erfüllt ist, ein zulässiges Ideal in A bildet, wenn $\mathfrak{b}$ ein derartiges Ideal in B ist. Wir behaupten: $a_l b \in T(\mathfrak{a})$, $ba_r \in T(\mathfrak{a})$, $b(ab'_r)_r \in T(\mathfrak{a})$ für alle $a \in \mathfrak{a}$ $b, b' \in B$ oder $\mathfrak{a} \subseteq T'(T(\mathfrak{a}))$.

Um die Richtigkeit der Behauptung zu beweisen, genügt es zu zeigen, daß $a_l b \in T(\mathfrak{a})$, $ba_r \in T(\mathfrak{a})$ für alle $a \in \mathfrak{a}$, $b \in B$. Wenn dies der Fall ist, so hat man: $b(ab'_r)_r = (ba_r)b' + (b_l a)_l b' - b(a_l b')$, wobei $ba_r \in T(\mathfrak{a})$, also $(ba_r)b' \in T(\mathfrak{a})$, weil $T(\mathfrak{a})$ zulässig ist. Aus $b_l a \in \mathfrak{a}$ ($\mathfrak{a}$ ist zulässig) folgt $(b_l a)_l b' \in T(\mathfrak{a})$ wegen der Annahme $a_l b \in T(\mathfrak{a})$ für alle $a \in \mathfrak{a}$, $b \in B$. Und da $a_l b' \in T(\mathfrak{a})$ hat man auch $b(a_l b') \in T(\mathfrak{a})$. Damit ist gezeigt, daß $b(ab'_r)_r \in T(\mathfrak{a})$ für alle $a \in \mathfrak{a}$, $b, b' \in B$.

Sei jetzt $a \in \mathfrak{a}$ und $b \in B$. Wir betrachten erst das Element $a_l b$. Man hat $(a_l b)_l a' = a(b_l a') + a(ba'_r)_r - (ab_r)a'$; $a(b_l a') \in \mathfrak{a}$, da $\mathfrak{a}$ Ideal in A ist, $a(ba'_r)_r \in \mathfrak{a}$, weil $\mathfrak{a}$ zulässig ist in A und $(ab_r)a' \in \mathfrak{a}$, da $\mathfrak{a}$ ein zulässiges Ideal in A ist. Also hat man $(a_l b)_l a' \in \mathfrak{a}$ *für alle* $a' \in A$. Und $a'(a_l b)_r = (a'a)b_r - a'(ab_r) \in \mathfrak{a}$ *für alle* $a' \in A$, da $\mathfrak{a}$ ein zulässiges Ideal in A ist. Wegen $a''((a_l b)a'_r)_r = a''(a_l(ba'_r))_r$ und $a''(a_l(ba'_r))_r \in \mathfrak{a}$ (soeben bewiesen), hat man auch $a''((a_l b)a'_r)_r \in \mathfrak{a}$ *für alle* $a'', a' \in A$. Nach (7) folgt hieraus $a_l b \in T(\mathfrak{a})$.

Für das Element ba_r beweist man analog, daß es zu $T(\mathfrak{a})$ gehört nach der Definition von $T(\mathfrak{a})$ in (7). Damit ist unsere Behauptung bewiesen. Wir fassen die Ergebnisse zusammen in

Satz 1. *Ist $R = A \dotplus B$ eine Szépsche Erweiterung von A und B, dann kann man zu jedem zulässigen Ideal $\mathfrak{a}$ in A ein Ideal $T(\mathfrak{a})$ in B konstruieren nach* (7), *und $T(\mathfrak{a})$ ist ein zulässiges Ideal in B. Umgekehrt gehört zu jedem zulässigen Ideal $\mathfrak{b}$ in B ein zulässiges Ideal $T'(\mathfrak{b})$ in A nach* (8). *Dabei hat man:* $\mathfrak{a} \subseteqq T'(T(\mathfrak{a}))$ *und* $\mathfrak{b} \subseteqq T(T'(\mathfrak{b}))$.

Bemerkung. Anfangend mit einem zulässigen Ideal $\mathfrak{a}$ in A kann man also eine nicht absteigende Kette von zulässigen Idealen in A bilden:

$$\mathfrak{a} \subseteqq T'(T(\mathfrak{a})) \subseteqq T'(T(T'(T(\mathfrak{a})))) \subseteqq \dots .$$

In B hat man dann die nicht absteigende Kette:

$$T(\mathfrak{a}) \subseteqq T(T'(T(\mathfrak{a}))) \subseteqq T(T'(T(T'(T(\mathfrak{a}))))) \subseteqq \dots .$$

Es scheint uns eine interessante Untersuchung zu sein, wenn man den zulässigen Idealen z. B. in A eine Endlichkeitsbedingung auflegt, in der Art, daß die soeben betrachtete Kette von zulässigen Idealen in A im Endlichen abbricht. Dies hat zur Folge, daß auch die korrespondierende Kette von zulässigen Idealen in B im Endlichen abbricht. Gibt es vielleicht Szépsche Erweiterungsringe $R = A \dotplus B$ wofür dies für *jedes* zulässige Ideal in A der Fall ist?

3. Die zulässigen Ideale in A bzw. B der Struktur $T'(\mathfrak{b})$ bzw. $T(\mathfrak{a})$, wo $\mathfrak{a}$ bzw. $\mathfrak{b}$ zulässige Ideale in A bzw. B sind, kann man benutzen bei den Untersuchungen über die Ideale in $R = A \dotplus B$.

Es sei wieder R eine Szépsche Erweiterung von A und B und $\mathfrak{a}$ ein gegebenes Ideal in A, das zulässig ist. Wir betrachten die Menge $K(\mathfrak{a})$ von Elementen $b \in B$, für die

(9) $$b_l a \in \mathfrak{a}, \quad ab_r \in \mathfrak{a}$$

für alle $a \in A$ erfüllt ist.

Es ist klar, daß $T(\mathfrak{a}) \subseteqq K(\mathfrak{a})$. Aus dem Beweis, daß $T(\mathfrak{a})$ ein Ideal in B ist, ergibt sich unmittelbar, daß auch $K(\mathfrak{a})$ ein Ideal in B ist. $K(\mathfrak{a})$ ist aber im allgemeinen nicht zulässig, wie SZENDREI mit einem Beispiel für $\mathfrak{a} = 0$ gezeigt hat. Man kann $K(\mathfrak{a})$ auch als den Kern eines Homomorphismus von B auf einen Ring B^* bekommen wie folgt: Man kann die Funktionen $b_l a$, ab_r als Operatorprodukte auffassen. In diesem Sinne ist A gleichzeitig ein Links- und Rechtsoperatorenbereich von B. Da $\mathfrak{a}$ ein Ideal in A ist, kann man den Restklassenring $A/\mathfrak{a}$ bilden und $A/\mathfrak{a}$ als Links- und Rechtsoperatorenbereich sehen nach der Definition:

(10) $$b_l(a+\mathfrak{a}) = b_l a + \mathfrak{a}, \quad (a+\mathfrak{a})b_r = ab_r + \mathfrak{a}.$$

Diese Definition ist unabhängig von der speziellen Wahl der Representanten in $a+\mathfrak{a}$. Ist $a' \equiv a(\mathfrak{a})$, dann gilt $b_l(a'+\mathfrak{a}) = b_l a' + \mathfrak{a}$, aber $b_l(a'-a) = b_l a' - b_l a \in \mathfrak{a}$, weil $a'-a \in \mathfrak{a}$, also $b_l a' + \mathfrak{a} = b_l a + \mathfrak{a}$, und ebenso $a'b_r + \mathfrak{a} = ab_r + \mathfrak{a}$. Aus (10) folgt, daß jedes Element b von B eine Doppelabbildung: $a+\mathfrak{a} \to b_l(a+\mathfrak{a})$, $a+\mathfrak{a} \to (a+\mathfrak{a})b_r$ von $A/\mathfrak{a}$ in sich induziert. Diese durch b induzierte Doppelabbildung von $A/\mathfrak{a}$ in sich bezeichnen wir mit b^* und die Menge von b^* mit B^*. Aus (3) folgt, daß jede

Doppelabbildung b^* ein Doppelendomorphismus von $A/\mathfrak{a}$ ist. Für die Addition und Multiplikation dieser Doppelabbildungen definieren wir:

$$b^* + b'^* = (b+b')^*, \quad b^*b'^* = (bb')^*. \tag{11}$$

Man kann nun beweisen, daß B^* mit dieser Addition und Multiplikation einen zu B homomorphen Ring bildet: d. h. $B \sim B^*(b \to b^*)$. Also besteht die homomorphe Abbildung einfach darin, daß man jedem Ringelement b die durch b induzierte Doppelabbildung b^* zuordnet. Der Kern dieses Homomorphismus ist die Menge der b mit $b_l(a+\mathfrak{a}) = \mathfrak{a}, (a+\mathfrak{a})b_r = \mathfrak{a}$ für alle $a \in A$ oder $b_l a \in \mathfrak{a}, ab_r \in \mathfrak{a}$ für alle $a \in A$, d. h. der Kern ist $K(\mathfrak{a})$, nach der Definition in (9).

Ebenso ist die Menge $K'(\mathfrak{b})$ der Elemente $a \in A$, für die $a_l b \in \mathfrak{b}, ba_r \in \mathfrak{b}$ für alle $b \in B$ erfüllt ist, ein Ideal in A, wenn $\mathfrak{b}$ ein zulässiges Ideal in B ist. Man hat: $T'(\mathfrak{b}) \subseteqq K'(\mathfrak{b})$.

Sind S bzw. T Teilmengen von A bzw. B, dann verstehen wir unter (S, T) die Gesamtheit derjenigen Elemente (a, b) aus R, für die $a \in S, b \in T$ gelten. Wir beweisen den

Satz 2. *$R = A \dotplus B$ sei eine Szépsche Erweiterung von A und B; $\mathfrak{a}$ bzw. $\mathfrak{b}$ sind zulässige Ideale in A bzw. B. Dann ist $(\mathfrak{a}, \mathfrak{b})$ ein Ideal in R dann und nur dann wenn: $\mathfrak{b} \subseteqq K(\mathfrak{a})$ und $\mathfrak{a} \subseteqq K'(\mathfrak{b})$.*

Beweis. Sei $a \in A$ und $b \in B$ und $(a_i, b_j) \in (\mathfrak{a}, \mathfrak{b})$. Offenbar gilt in R: $(a_i, b_j) - (a_i', b_j') = (a_i - a_i', b_j - b_j') \in (\mathfrak{a}, \mathfrak{b})$, wenn $(a_i', b_j') \in (\mathfrak{a}, \mathfrak{b})$ da $\mathfrak{a}$ bzw. $\mathfrak{b}$ Ideale in A bzw. B sind. Weiter hat man: $(a, 0)(a_i, b_j) = (aa_i + ab_{j_r}, a_l b_j) \in (\mathfrak{a}, \mathfrak{b})$ dann und nur dann wenn $ab_{j_r} \in \mathfrak{a}$, weil aa_i bzw. $a_l b_j$ zu $\mathfrak{a}$ bzw. $\mathfrak{b}$ gehören; $(0, b)(a_i, b_j) = (b_l a_i, ba_{i_r} + bb_j) \in (\mathfrak{a}, \mathfrak{b})$ dann und nur dann wenn $ba_{i_r} \in \mathfrak{b}$, weil $b_l a_i$ bzw. bb_j zu $\mathfrak{a}$ bzw. $\mathfrak{b}$ gehören; ebenso $(a_i, b_j)(a, 0) \in (\mathfrak{a}, \mathfrak{b}) \leftrightarrow b_{j_l} a \in \mathfrak{a}$, und $(a_i, b_j)(0, b) \in (\mathfrak{a}, \mathfrak{b}) \leftrightarrow a_{i_l} b \in \mathfrak{b}$. Die notwendige und hinreichende Bedingung damit für (a_i, b_j) beliebig in $(\mathfrak{a}, \mathfrak{b})$ auch $(a, b)(a_i, b_j)$ und $(a_i, b_j)(a, b)$ zu $(\mathfrak{a}, \mathfrak{b})$ gehören ist also: $b_{j_l} a, ab_{j_r} \in \mathfrak{a}$ für alle $a \in A, b_j \in \mathfrak{b}$ und $a_{i_l} b, ba_{i_r} \in \mathfrak{b}$ für alle $b \in B, a_i \in \mathfrak{a}$. Man kann auch sagen: $(\mathfrak{a}, \mathfrak{b})$ ist ein Ideal in R dann und nur dann, falls $b_j \in K(\mathfrak{a})$ für alle $b_j \in \mathfrak{b}$ oder $\mathfrak{b} \subseteqq K(\mathfrak{a})$ und $a_i \in K'(\mathfrak{b})$ für alle $a_i \in \mathfrak{a}$, oder $\mathfrak{a} \subseteqq K'(\mathfrak{b})$. Damit ist der Satz bewiesen.

Wir behaupten nun: aus $\mathfrak{a} \subseteqq K'(\mathfrak{b})$ folgt $\mathfrak{a} \subseteqq T'(\mathfrak{b})$. Denn, wenn $b, b' \in B$, $a \in \mathfrak{a} \subseteqq K'(\mathfrak{b})$, so brauchen wir nur noch zu zeigen, daß $b(ab_r')_r \in \mathfrak{b}$. Man hat aber: $b(ab_r')_r = (ba_r)b' + (b_l a)_l b' - b(a_l b')$. Da $ba_r \in \mathfrak{b}$ so folgt $(ba_r)b' \in \mathfrak{b}$; da $a_l b' \in \mathfrak{b}$, so ist auch $b(a_l b') \in \mathfrak{b}$; und wegen $b_l a \in \mathfrak{a}$ hat man $(b_l a)_l b' \in \mathfrak{b}$. Hieraus folgt $b(ab_r')_r \in \mathfrak{b}$ für alle $b, b' \in B$, $a \in \mathfrak{a}$, weshalb $a \in T'(\mathfrak{b})$. Damit ist gezeigt: Wenn $\mathfrak{a}$ ein zulässiges Ideal in A ist und $\mathfrak{a} \subseteqq K'(\mathfrak{b})$, so gilt auch $\mathfrak{a} \subseteqq T'(\mathfrak{b})$. Ebenso: $\mathfrak{b} \subseteqq K(\mathfrak{a})$ impliziert $\mathfrak{b} \subseteqq T(\mathfrak{a})$.

Da $T'(\mathfrak{b}) \subseteqq K'(\mathfrak{b})$ auch ein zulässiges Ideal in A ist, so kann man sagen: *$T'(\mathfrak{b})$ ist das größte, in $K'(\mathfrak{b})$ enthaltene zulässige Ideal in A, und ebenso ist $T(\mathfrak{a})$ das größte, in $K(\mathfrak{a})$ enthaltene zulässige Ideal in B.*

Man hat auch: $T'(\mathfrak{b})(T(\mathfrak{a}))$ und $K'(\mathfrak{b})(K(\mathfrak{a}))$ enthalten *dieselbe* zulässige Ideale von $A(B)$. Den Satz 2 kann man nun ersetzen durch:

Satz 2a. *$R = A \dotplus B$ ist eine Szépsche Erweiterung von A und B; $\mathfrak{a}$ bzw. $\mathfrak{b}$ sind zulässige Ideale in A bzw. B. Dann ist $(\mathfrak{a}, \mathfrak{b})$ dann und nur dann ein Ideal in R wenn: $\mathfrak{a} \subseteqq T'(\mathfrak{b})$ und $\mathfrak{b} \subseteqq T(\mathfrak{a})$.*

Einige Spezialfälle des Satzes 2a sind:

a.) $(\mathfrak{a}, T(\mathfrak{a}))$ ist ein Ideal in R dann und nur dann wenn $\mathfrak{a} \subseteqq T'(T(\mathfrak{a}))$ und $T(\mathfrak{a}) \subseteqq T(\mathfrak{a})$. Da $\mathfrak{a} \subseteqq T'(T(\mathfrak{a}))$ immer gilt, so ist $(\mathfrak{a}, T(\mathfrak{a}))$ ein Ideal in R. Ebenso ist $(T'(\mathfrak{b}), \mathfrak{b})$ ein ideal in R.

Wenn $(\mathfrak{a}^*, \mathfrak{b}^*)$ ein Ideal in R ist mit $\mathfrak{a}^* \subseteqq \mathfrak{a}$, so hat man $\mathfrak{b}^* \subseteqq T(\mathfrak{a})$. Aus der Definition von $T(\mathfrak{a})$ in (7) ist unmittelbar klar, daß aus $\mathfrak{a}^* \subseteqq \mathfrak{a}$ folgt $T(\mathfrak{a}^*) \subseteqq T(\mathfrak{a})$. Also $\mathfrak{b}^* \subseteqq T(\mathfrak{a}^*) \subseteqq T(\mathfrak{a})$. Bei festem zulässigen Ideal $\mathfrak{a}$ in A gilt also für jedes Ideal $(\mathfrak{a}, \mathfrak{b}^*)$ in R, daß $\mathfrak{b}^* \subseteqq T(\mathfrak{a})$. Ebenso hat man bei festem zulässigen Ideal $\mathfrak{b}$ in B für jedes Ideal $(\mathfrak{a}^*, \mathfrak{b})$ in R, daß $\mathfrak{a}^* \subseteqq T'(\mathfrak{b})$.

b.) $(T'(\mathfrak{b}), T(\mathfrak{a}))$ ist ein Ideal in R dann und nur dann wenn $T'(\mathfrak{b}) \subseteqq T'(T(\mathfrak{a}))$, $T(\mathfrak{a}) \subseteqq T(T'(\mathfrak{b}))$. Dies ist immer der Fall wenn $(\mathfrak{a}, \mathfrak{b})$ Ideal in R ist, denn aus $\mathfrak{b} \subseteqq T(\mathfrak{a})$ folgt $T'(\mathfrak{b}) \subseteqq T'(T(\mathfrak{a}))$ und aus $\mathfrak{a} \subseteqq T'(\mathfrak{b})$ folgt $T(\mathfrak{a}) \subseteqq T(T'(\mathfrak{b}))$.

c.) $(\mathfrak{a}, 0)$ ist ein Ideal in R dann und nur dann wenn $\mathfrak{a} \subseteqq T'(0)$. Da $(T'(0), 0)$ ein Ideal in R ist (sieh a), so folgt, daß jedes Ideal $(\mathfrak{a}, 0)$ in R eine Teilmenge des Ideals $(T'(0), 0)$ ist oder $(T'(0), 0)$ ist das größte in $(A, 0)$ enthaltene Ideal von R. Ebenso ist $(0, T(0))$ das größte in $(0, B)$ enthaltene Ideal von R. Weiter ist $(T'(0), T(0))$ ein Ideal in R, da $(0, 0)$ ein Ideal (das Nullideal) in R ist (sieh b).) Wir bemerken, daß SZENDREI [3] das in c.) gefundene schon in seinem Satz 2 bewiesen hat.

d.) $(0, T(\mathfrak{a}))$ ist ein Ideal in R dann und nur dann wenn $T(\mathfrak{a}) \subseteqq T(0)$ oder $T(\mathfrak{a}) = T(0)$, da $T(0) \subseteqq T(\mathfrak{a})$ für jedes zulässige Ideal $\mathfrak{a}$ in A. Ebenso ist $(T'(\mathfrak{b}), 0)$ ein Ideal in R dann und nur dann wenn $T'(\mathfrak{b}) = T'(0)$.

Definition 2. Unter einem *echten* Ideal S im Ring R verstehen wir ein Ideal $S \neq 0$ und $\neq R$.

Es sei $R = A \dotplus B$ eine Szépsche Erweiterung von A und B. Ist $(\mathfrak{a}, \mathfrak{b})$ ein echtes Ideal in R, wobei $\mathfrak{a}$ bzw. $\mathfrak{b}$ Ideale in A bzw. B sind, dann sind $\mathfrak{a}$ bzw. $\mathfrak{b}$ auch zulässig in A bzw. B. Dabei ist $\mathfrak{a} \neq A$ oder $\mathfrak{b} \neq B$ oder beide und $\mathfrak{a} \neq 0$ oder $\mathfrak{b} \neq 0$ oder beide. Also enthält wenigstens der eine von A und B ein echtes zulässiges Ideal. Umgekehrt, sei $\mathfrak{a}$ ein echtes zulässiges Ideal in A. Dann ist $(\mathfrak{a}, T(\mathfrak{a}))$ ein Ideal in R und $(\mathfrak{a}, T(\mathfrak{a})) \neq (0, 0)$ denn $\mathfrak{a} \neq 0$, $(\mathfrak{a}, T(\mathfrak{a})) \neq (A, B)$ denn $\mathfrak{a} \neq A$. Also ist $(\mathfrak{a}, T(\mathfrak{a}))$ ein echtes Ideal in R.

Eine notwendige und hinreichende Bedingung dafür, daß R ein echtes Ideal der Form $(\mathfrak{a}, \mathfrak{b})$ enthält, wo $\mathfrak{a}$ bzw. $\mathfrak{b}$ Ideale in A bzw. B sind, ist, daß wenigstens der eine von A und B ein echtes zulässiges Ideal enthält.

Satz 3. *$R = A \dotplus B$ ist eine Szépsche Erweiterung von A und B, in der weder A noch B ein Ideal ist und weder A noch B zulässige Unterringe enthalten. Ist I ein echtes Ideal in R, dann ist I^+ subdirekte Summe von A^+ und B^+.*

Beweis. Ist I ein Ideal von $R = A \dotplus B$ bestehend aus den Elementen $a_1 + b_1$, $a_2 + b_2, \ldots$ $(a_i \in A, b_i \in B, i = 1, 2, \ldots)$, so bilden die a_i bzw. b_i Unterringe A' bzw. B' von R. ([4], Satz 7). Wenn a ein beliebiges Element von A ist und $(a_i, b_i) \in I$, dann $(a, 0)(a_i, b_i) = (aa_i + ab_{i_l}, a_l b_i) \in I$, und es folgt, daß $a_l b_i \in B'$. Ähnlich folgt aus $(a_i, b_i)(a, 0) = (a_i a + b_{i_l} a, b_i a_r) \in I$ daß $b_i a_r \in B'$. Auf demselben Weg können wir $a_i b_r$, $b_l a_i \in A'$ $(b \in B)$ zeigen. Also sind A' bzw. B' zulässige Unterringe von A bzw. B nach Definition 1. Da A noch B einen echten zulässigen Unterring enthalten, hat man: $A' = 0$ oder $A' = A$ und ebenso $B' = 0$ oder $B' = B$. Man hat also 4 Fälle zu unterscheiden:

a.) $A'=0$ und $B'=0$. In diesem Falle ist $I=0$, also kein echtes Ideal in R. Widerspruch.

b.) $A'=0$ und $B'=B$. Dann ist $I\subseteq(0, B)$ und I ist ein Ideal in $(0, B)$. Wegen $(a, 0)(0, b_i) = (ab_{i_r}, a_l b_i)\in I$ für jedes $a\in A$ hat man $ab_{i_r}=0$ und $(0, a_l b_i)\in I$ für jedes $a\in A$. Ebenso ist $(0, b_i a_r)\in I$ für jedes $a\in A$. I ist also ein zulässiges Ideal in B (oder $(0, B)$). Da nun B keine zulässige Ideale enthält hat man $I=(0, 0)$ oder $I=(0, B)$. Wegen I ist ein echtes Ideal in R und B ist kein Ideal in R, hat man wieder einen Widerspruch.

c.) $A'=A$ und $B'=0$. Dann ist $I\subseteq(A, 0)$ und auf dieselbe Weise wie unter b. zeigt man, daß I ein zulässiges Ideal in A ist. Daraus folgt dann, daß $I=(0, 0)$ oder $I=(A, 0)$, was nicht der Fall ist, also ist auch diese Möglichkeit ausgeschlossen.

d.) $A'=A$ und $B'=B$. Dann ist I^+ ein Untermodul von $R^+ = (A, 0)^+\oplus(0, B)^+$ und durch die Zuordnung $(a_i, b_i)\to(a_i, 0)$ bzw. $(a_i, b_i)\to(0, b_i)$ wird ein Homomorphismus von I^+ auf $(A', 0)$ bzw. $(0, B')$ definiert. Da aber $A'=A$ und $B'=B$, so ist I^+ subdirekte Summe der Untermoduln $(A, 0)$ und $(0, B)$. Damit ist der Satz bewiesen.

4. Wir beschäftigen uns nun mit denjenigen Ringen $R = A\dotplus B$, für welche

$$\mathfrak{a}=T'(T(\mathfrak{a})) \tag{12}$$

für jedes zulässige Ideal $\mathfrak{a}$ in A erfüllt ist und wobei jedes zulässige Ideal $\mathfrak{b}$ von B in der Form $\mathfrak{b}=T(\mathfrak{a})$ geschrieben werden kann.

Wir betrachten die Abbildung $T\colon \mathfrak{a}\to T(\mathfrak{a})$, die jedem zulässigen Ideal in A ein zulässiges Ideal in B zuordnet. Diese Abbildung ist eineindeutig. Es ist klar, daß wenn $\mathfrak{a}=\mathfrak{a}'$, dann auch $T(\mathfrak{a})=T(\mathfrak{a}')$ ist nach der Definition in (7). Umgekehrt, wenn $\mathfrak{b}=T(\mathfrak{a})=T(\mathfrak{a}')$, dann folgt $\mathfrak{a}=T'(T(\mathfrak{a}))=T'(\mathfrak{b})$ und $\mathfrak{a}'=T'(T(\mathfrak{a}'))=T'(\mathfrak{b})$, also $\mathfrak{a}=\mathfrak{a}'$. Für jedes zulässige Ideal $\mathfrak{b}$ in B gilt: aus $\mathfrak{b}=T(\mathfrak{a})$ folgt $T'(\mathfrak{b})=T'(T(\mathfrak{a}))=\mathfrak{a}$, also ist $\mathfrak{b}=T(T'(\mathfrak{b}))$ für jedes zulässige Ideal $\mathfrak{b}$ in B. Die Abbildung $T'\colon \mathfrak{b}\to T'(\mathfrak{b})$, die jedem zulässigen Ideal in B ein zulässiges Ideal in A zuordnet, ist also die inverse von T: $T'(\mathfrak{b})\to T(T'(\mathfrak{b}))=\mathfrak{b}$.

Da aus $\mathfrak{a}\subseteq\mathfrak{a}'$ die Relation $T(\mathfrak{a})\subseteq T(\mathfrak{a}')$ folgt, ist die Abbildung T ordnungbehaltend. Der Durchschnitt $\mathfrak{a}\cap\mathfrak{a}'$ von zwei zulässigen Idealen $\mathfrak{a}, \mathfrak{a}'$ in A ist wieder ein zulässiges Ideal in A. Man kann also $T(\mathfrak{a}\cap\mathfrak{a}')$ bilden. Aus der Definition in (7) folgt, daß $T(\mathfrak{a}\cap\mathfrak{a}') = T(\mathfrak{a})\cap T(\mathfrak{a}')$. Auch die Summe $\mathfrak{a}+\mathfrak{a}'$ von zwei zulässigen Idealen $\mathfrak{a}, \mathfrak{a}'$ in A ist wieder ein derartiges Ideal in A. Aus $\mathfrak{a}\subseteq\mathfrak{a}+\mathfrak{a}'$ und $\mathfrak{a}'\subseteq\mathfrak{a}+\mathfrak{a}'$ folgt $T(\mathfrak{a})\subseteq T(\mathfrak{a}+\mathfrak{a}')$ und $T(\mathfrak{a}')\subseteq T(\mathfrak{a}+\mathfrak{a}')$, also $T(\mathfrak{a})+T(\mathfrak{a}')\subseteq T(\mathfrak{a}+\mathfrak{a}')$. Aus $T(\mathfrak{a})\subseteq T(\mathfrak{a})+T(\mathfrak{a}')$ folgt $T'(T(\mathfrak{a}))\subseteq T'(T(\mathfrak{a})+T(\mathfrak{a}'))$ oder $\mathfrak{a}\subseteq T'(T(\mathfrak{a})+T(\mathfrak{a}'))$. Ebenso $\mathfrak{a}'\subseteq T'(T(\mathfrak{a})+T(\mathfrak{a}'))$. Wegen $T(\mathfrak{a})+T(\mathfrak{a}')\subseteq T(\mathfrak{a}+\mathfrak{a}')$ hat man $T'(T(\mathfrak{a})+ T(\mathfrak{a}'))\subseteq\mathfrak{a}+\mathfrak{a}'$. Setzen wir $T(\mathfrak{a})+T(\mathfrak{a}') = T(\mathfrak{b})$, so ist $T'(T(\mathfrak{b})) = \mathfrak{a}+\mathfrak{a}'$ oder $\mathfrak{b} = \mathfrak{a}+\mathfrak{a}'$, und $T(\mathfrak{b}) = T(\mathfrak{a})+T(\mathfrak{a}') = T(\mathfrak{a}+\mathfrak{a}')$.

Wir betrachten die Menge aller zulässigen Ideale $\mathfrak{a}$ in A. Mit der Relation der Inklusion ist diese Menge eine teilweise geordnete Menge. Es ist ein Verband in bezug auf die Operationen $\cap$ und $+$ wobei $\mathfrak{a}_1\cap\mathfrak{a}_2$ der Durchschnitt und $\mathfrak{a}_1+\mathfrak{a}_2$ die Summe der Idealen $\mathfrak{a}_1$ und $\mathfrak{a}_2$ bedeutet. Dieser Verband wird bezeichnet mit $L(A)$. Das Nullideal ist das kleinste und A das größte Element von $L(A)$. Aus unseren Betrachtungen folgt nun:

Satz 4. *Wenn die Ideale von $L(A)$ der Bedingung* (12) *genügen und jedes zulässige Ideal* $\mathfrak{b}$ *in* B *das Bild eines zulässigen Ideals* $\mathfrak{a}$ *in* A *unter die Abbildung* $T: \mathfrak{a} \to T(\mathfrak{a})$ *von* $L(A)$ *auf* $L(B)$ *ist, dann ist die Abbildung* $T: \mathfrak{a} \to T(\mathfrak{a})$ *ein Verbandsisomorphismus von* $L(A)$ *auf* $L(B)$.

Man kann auch jedem zulässigen Ideal $\mathfrak{a}$ in A das eindeutig bestimmte Ideal $(\mathfrak{a}, T(\mathfrak{a}))$ von R zuordnen. Aus $\mathfrak{a} \subseteq \mathfrak{a}'$ folgt $T(\mathfrak{a}) \subseteq T(\mathfrak{a}')$, also $(\mathfrak{a}, T(\mathfrak{a})) \subseteq (\mathfrak{a}', T(\mathfrak{a}'))$. Sind $\mathfrak{a}$ und $\mathfrak{a}'$ Elemente von $L(A)$ so ist $(\mathfrak{a}, T(\mathfrak{a})) \cap (\mathfrak{a}', T(\mathfrak{a}')) = (\mathfrak{a} \cap \mathfrak{a}', T(\mathfrak{a}) \cap \cap T(\mathfrak{a}')) = (\mathfrak{a} \cap \mathfrak{a}', T(\mathfrak{a} \cap \mathfrak{a}'))$ und $(\mathfrak{a}, T(\mathfrak{a})) \cup (\mathfrak{a}', T(\mathfrak{a}')) = (\mathfrak{a}, T(\mathfrak{a})) + (\mathfrak{a}', T(\mathfrak{a}')) = = (\mathfrak{a} + \mathfrak{a}', T(\mathfrak{a}) + T(\mathfrak{a}')) = (\mathfrak{a} + \mathfrak{a}', T(\mathfrak{a} + \mathfrak{a}'))$. Es folgt:

Satz 5. *Unter denselben Voraussetzungen wie in Satz* 4 *ist die Abbildung* $\mathfrak{a} \to (\mathfrak{a}, T(\mathfrak{a}))$ *ein Verbandsisomorphismus von* $L(A)$ *auf den Verband aller Ideale der Form* $(\mathfrak{a}, T(\mathfrak{a}))$ *in* R.

5. Es sei $R = A \dotplus B$ eine Szépsche Erweiterung von A und B. (Die Bedingung (12) setzen wir nicht voraus.) Die Ideale $\mathfrak{a}$ bzw. $\mathfrak{b}$ sind zulässige Ideale in A bzw. B, und $(\mathfrak{a}, \mathfrak{b})$ ein Ideal in R. Wir wollen nun den Restklassenring $R/(\mathfrak{a}, \mathfrak{b})$ untersuchen. Wir können eine Szépsche Erweiterung von $A/\mathfrak{a}$ und $B/\mathfrak{b}$ konstruieren, wenn wir definieren:

$$(a_1 + \mathfrak{a}, b_1 + \mathfrak{b}) + (a_2 + \mathfrak{a}, b_2 + \mathfrak{b}) = (a_1 + a_2 + \mathfrak{a}, b_1 + b_2 + \mathfrak{b}),$$

$$(a_1 + \mathfrak{a}, b_1 + \mathfrak{b})(a_2 + \mathfrak{a}, b_2 + \mathfrak{b}) =$$

$$= \big(a_1 a_2 + (b_1 + \mathfrak{b})_l (a_2 + \mathfrak{a}) + (a_1 + \mathfrak{a})(b_2 + \mathfrak{b})_r,\ (a_1 + \mathfrak{a})_l (b_2 + \mathfrak{b}) + (b_1 + \mathfrak{b})(a_2 + \mathfrak{a})_r + b_1 b_2\big),$$

wobei

$$(b_1 + \mathfrak{b})_l (a_2 + \mathfrak{a}) = b_{1_l} a_2 + \mathfrak{a}, \quad (a_1 + \mathfrak{a})(b_2 + \mathfrak{b})_r = a_1 b_{2_r} + \mathfrak{a},$$

$$(a_1 + \mathfrak{a})_l (b_2 + \mathfrak{b}) = a_{1_l} b_2 + \mathfrak{b}, \quad (b_1 + \mathfrak{b})(a_2 + \mathfrak{a})_r = b_1 a_{2_r} + \mathfrak{b}.$$

Aus der Tatsache, daß $\mathfrak{a}$ bzw. $\mathfrak{b}$ zulässige Ideale in A bzw. B sind, kann man leicht beweisen, daß diese Definitionen der Funktionen unabhängig von der speziellen Wahl der Representanten von $A/\mathfrak{a}$ bzw. $B/\mathfrak{b}$ sind. Da $R = A \dotplus B$ eine Szépsche Erweiterung von A und B ist, so folgt daß die Funktionen $b_l a$, $ab_r (\in A)$ und $a_l b$, $ba_r (\in B)$ den Bedingungen (1)—(6) unterworfen sind. Deswegen genügen auch die oben eingeführten Funktionen in $A/\mathfrak{a}$ und $B/\mathfrak{b}$ den Bedingungen (1)—(6). Damit haben wir eine Szépsche Erweiterung R^* von $A/\mathfrak{a}$ und $B/\mathfrak{b}$ gefunden und es ist klar, daß die Abbildung $(a, b) \to (a + \mathfrak{a}, b + \mathfrak{b})$ ein Homomorphismus von $R = A \dotplus B$ auf R^* ist. Der Kern dieses Homomorphismus ist das Ideal $(\mathfrak{a}, \mathfrak{b})$. Wir haben bewiesen den

Satz 6. *Ist* $R = A \dotplus B$ *eine Szépsche Erweiterung von* A *und* B *und* $(\mathfrak{a}, \mathfrak{b})$ *ein Ideal in* R, *wobei* $\mathfrak{a}$ *bzw.* $\mathfrak{b}$ *zulässige Ideale in* A *bzw.* B *sind, dann ist* $R/(\mathfrak{a}, \mathfrak{b}) \cong \cong A/\mathfrak{a} \dotplus B/\mathfrak{b}$.

Bemerkung. Im spezialen Fall $\mathfrak{a} = T'(0)$ und $\mathfrak{b} = T(0)$, hat man Satz 1 meiner Arbeit [1] gefunden.

Ist insbesondere $A = K'(\mathfrak{b})$ und $B = K(\mathfrak{a})$ für das Ideal $(\mathfrak{a}, \mathfrak{b})$ in R, dann gilt: $R/(\mathfrak{a}, \mathfrak{b}) \cong A/\mathfrak{a} \oplus B/\mathfrak{b}$ (direkte Summe). Denn $A = K'(\mathfrak{b})$ bedeutet, daß $b_l a, ab_r \in \mathfrak{a}$ für alle $a \in A$ und $b \in B$ erfüllt ist, und das hat zur Folge, daß $(b + \mathfrak{b})_l (a + \mathfrak{a}) = \mathfrak{a}$, $(a + \mathfrak{b})(b + \mathfrak{b})_r = \mathfrak{a}$, für alle $b \in B$ und $a \in A$ gilt. Ebenso folgt aus $B = K(\mathfrak{a})$, daß

$(a+\mathfrak{a})_l(b+\mathfrak{b}) = \mathfrak{b}$ und $(b+\mathfrak{b})(a+\mathfrak{a})_r = \mathfrak{b}$ für alle $a \in A$ und $b \in B$ erfüllt ist. Für die Multiplikation in $R^* = R/(\mathfrak{a}, \mathfrak{b})$ hat man dann: $(a_1+\mathfrak{a}, b_1+\mathfrak{b})(a_2+\mathfrak{a}, b_2+\mathfrak{b}) = (a_1a_2+\mathfrak{a}, b_1b_2+\mathfrak{b})$, woraus die Behauptung unmittelbar vorgeht.

Umgekehrt entsteht nun das Problem: Es seien A und B gegebene Ringe und $A/\mathfrak{a} \dotplus B/\mathfrak{b}$ ist eine gegebene Szépsche Erweiterung von $A/\mathfrak{a}$ und $B/\mathfrak{b}$. Dabei sind $\mathfrak{a}$ bzw. $\mathfrak{b}$ Ideale in A bzw. B. Kann man eine Szépsche Erweiterung $R = A \dotplus B$ von A und B konstruieren, derart, daß $\mathfrak{a}^+ \oplus \mathfrak{b}^+$ Ideal in R ist und $R/(\mathfrak{a}^+ \oplus \mathfrak{b}^+) \cong A/\mathfrak{a} \dotplus B/\mathfrak{b}$ ist?

Eine teilweise Lösung dieses Problems wird gegeben in

Satz 7. *A und B sind gegebene Ringe und $R' = A/\mathfrak{a} \dotplus B/\mathfrak{b}$ ist eine Szépsche Erweiterung von $A/\mathfrak{a}$ und $B/\mathfrak{b}$, $\mathfrak{a}$ bzw. $\mathfrak{b}$ sind Ideale in A bzw. B. Eine hinreichende Bedingung, damit es eine Szépsche Erweiterung R von A und B gibt, in der $\mathfrak{a}^+ \oplus \mathfrak{b}^+$ ein Ideal ist und derart, daß $R/\mathfrak{a}^+ \oplus \mathfrak{b}^+ \cong A/\mathfrak{a} \dotplus B/\mathfrak{b}$ ist, daß die Restklassenringe $A/\mathfrak{a}$ bzw. $B/\mathfrak{b}$ je einen Repräsentantensystem enthalten, das einen Ring bildet.*

Beweis. Die Elemente von $R' = A/\mathfrak{a} \dotplus B/\mathfrak{b}$ sind die Paare $(a+\mathfrak{a}, b+\mathfrak{b})$ $(a+\mathfrak{a} \in A/\mathfrak{a}, b+\mathfrak{b} \in B/\mathfrak{b})$ mit der Addition

$$(a+\mathfrak{a}, b+\mathfrak{b})+(a'+\mathfrak{a}, b'+\mathfrak{b}) = (a+a'+\mathfrak{b}, b+b'+\mathfrak{b})$$

und der Multiplikation

$$(a+\mathfrak{a}, b+\mathfrak{b})(a'+\mathfrak{a}, b'+\mathfrak{b}) =$$

$$=\big(aa'+(b+\mathfrak{b})_l(a'+\mathfrak{a})+(a+\mathfrak{a})(b'+\mathfrak{a})_r, (a+\mathfrak{a})_l(b'+\mathfrak{b})+(b+\mathfrak{b})(a'+\mathfrak{a})_r+bb'\big),$$

wobei die Funktionen $(b+\mathfrak{b})_l(a'+\mathfrak{a})$, $(a+\mathfrak{a})(b'+\mathfrak{b})_r$ $(\in A/\mathfrak{a})$ und $(a+\mathfrak{a})_l(b'+\mathfrak{b})$, $(b+\mathfrak{b})(a'+\mathfrak{a})_r$ $(\in B/\mathfrak{b})$ den Bedingungen (D), (1)—(6) unterworfen sind. Es sei nun $f(a+\mathfrak{a})$ $(\in a+\mathfrak{a})$ bzw. $g(b+\mathfrak{b})$ $(\in b+\mathfrak{b})$ ein Repräsentantensystem in $A/\mathfrak{a}$ bzw. $B/\mathfrak{b}$, das einen Ring bildet; d. h.

$$f(a+\mathfrak{a})+f(a'+\mathfrak{a}) = f(a+a'+\mathfrak{a}); \; f(\mathfrak{a})=0;$$

$f(a+\mathfrak{a})f(a'+\mathfrak{a}) = f(aa'+\mathfrak{a})$ und ähnlich für $g(b+\mathfrak{b})$. Die Restklassen mod $\mathfrak{a}$ lassen sich also mit den sämtlichen Elementen eines Unterringes von A repräsentieren, d. h. A „zerfällt" in eine schlichte Summe des Ideals $\mathfrak{a}$ und des Unterringes der $f(a+\mathfrak{a})$. Ebenso zerfällt B in die Summe des Ideals $\mathfrak{b}$ und des Unterringes der $g(b+\mathfrak{b})$. Die „Funktionenvierer" b_la, $ab_r (\in A)$, a_lb, $ba_r (\in B)$ für die Szépsche Erweiterung $R = A \dotplus B$ definieren wir wie folgt:

$$b_la = f\big((b+\mathfrak{b})_l(a+\mathfrak{a})\big), \; ab_r = f\big((a+\mathfrak{a})(b+\mathfrak{b})_r\big),$$

$$a_lb = g\big((a+\mathfrak{a})_l(b+\mathfrak{b})\big), \; ba_r = g\big((b+\mathfrak{b})(a+\mathfrak{a})_r\big).$$

Ist $b^* \equiv b \pmod{\mathfrak{b}}$ und $a^* \equiv a \pmod{\mathfrak{a}}$, dann ist $b^*_la^* = f\big((b^*+\mathfrak{b})_l(a^*+\mathfrak{a})\big) = f\big((b+\mathfrak{b})_l(a+\mathfrak{a})\big) = b_la$ und ähnlich für ab_r, a_lb und ba_r. Wegen $f(\mathfrak{a})=0$ und $g(\mathfrak{b})=0$ genügen die Funktionen der Bedingung (D). Man hat nun $(b+b')_la = f\big((b+b'+\mathfrak{b})_l(a+\mathfrak{a})\big) = f\big((b+\mathfrak{b})_l(a+\mathfrak{a})+(b'+\mathfrak{b})_l(a+\mathfrak{a})\big) = f\big((b+\mathfrak{b})_l(a+\mathfrak{a})\big)+f\big((b'+\mathfrak{b})_l(a+\mathfrak{a})\big) = b_la+b'_la$ und wenn man benutzt, daß die Repräsentanten $f(a+\mathfrak{a})$ bzw. $g(b+\mathfrak{b})$ einen Ring in A bzw. B bilden, beweist man die übrigen Formeln von (1) und auch (2)—(6) auf dieselbe Weise. Damit hat man eine Szépsche

Erweiterung $R = A + B$ von A und B konstruiert. Es ist klar, daß die Abbildung $(a, b) \to (a + \mathfrak{a}, b + \mathfrak{b})$ ein Homomorphismus von R auf R' ist, dessen Kern das Ideal $(\mathfrak{a}, \mathfrak{b})$ in R ist. Also ist $R/(\mathfrak{a}, \mathfrak{b}) \cong R'$ und der Satz bewiesen.

Man kann noch bemerken, daß das Ideal $(\mathfrak{a}, \mathfrak{b})$ hier die direkte Summe der Ideale $(\mathfrak{a}, 0)$ und $(0, \mathfrak{b})$ in $(A, 0)$ und $(0, B)$ ist.

Literaturverzeichnis

[1] L. C. A. van Leeuwen, On ring extensions of Szép, *Nederl. Akad. Wet. Proc.*, ser. A, **67** (1964), 40–47.
[2] L. Rédei, *Algebra*. I (Leipzig, 1959).
[3] J. Szendrei, Über die Szépschen Ringerweiterungen, *Acta Sci. Math.*, **21** (1960), 166–172.
[4] J. Szép, Über eine neue Erweiterung von Ringen. I, *Acta Sci. Math.*, **19** (1958), 51–62.
[5] J. Szép, Über eine neue Erweiterung von Ringen. II, *Acta Sci. Math.*, **20** (1959), 202–214.

(Eingegangen am 18. April 1964)

Über die Existenz normaler Komplemente zu gewissen Hallgruppen

Von LUDWIG PROHASKA in Rostock (DDR)

1. G sei eine endliche Gruppe, e ihr Einselement. $N \subseteq| G$ bezeichne einen Normalteiler N von G. N heißt *normales Komplement* zu der Untergruppe U von G, wenn $G = UN$ mit $U \cap N = \{e\}$ und $N \subseteq| G$ ist. Hat in der Nebenklassenzerlegung

$$G = \sum_{r_i \in R} Ur_i$$

von G nach U das Repräsentantensystem R die Eigenschaft

$$u^{-1}Ru = R \quad \text{für jedes } u \in U,$$

so wird R *ausgezeichnetes Repräsentantensystem* genannt [3].

Mit Hilfe dieses Begriffs läßt sich ein bekannter Satz von BURNSIDE [1], S. 327, folgendermaßen aussprechen [3]:

Eine abelsche Sylowgruppe P von G hat genau dann ein ausgezeichnetes Repräsentantensystem in G, wenn sie in G ein normales Komplement besitzt.

KOCHENDÖRFFER [3] und ZAPPA [4] zeigten, daß hierin eine geringere Voraussetzung über P an die Stelle von „abelsche Sylowgruppe" treten kann. ZAPPA setzte dafür „nilpotente Hallgruppe" und wies in einer Anmerkung zu [4] darauf hin, daß noch eine weitere Abschwächung möglich ist. Dafür soll in der vorliegenden Arbeit ein Beweis gegeben werden, der sich stützt auf den

Satz (FROBENIUS [2]). *P sei eine p-Sylowgruppe von G. Für jedes Element $g \in G$ mit zu p teilerfremder Ordnung und für jede Untergruppe Q von G, deren Ordnung eine Potenz von p ist, möge gelten: wenn g im Normalisator $N_G(Q)$ von Q in G liegt, dann sogar im Zentralisator $Z_G(Q)$ von Q in G.*

Dann besitzt P in G ein normales Komplement.

2. π sei eine Primzahlmenge. Eine Gruppe wird *π-Gruppe* genannt, wenn jeder Primfaktor ihrer Ordnung in π enthalten ist.

Von ZAPPA [4] wurde angegeben das

Lemma. *Sei H eine Hallgruppe von G, die in G ein ausgezeichnetes Repräsentantensystem R besitzt. π bezeichne die Menge der Primteiler der Ordnung $|H|$ von H. Für jede Untergruppe $Q \subseteq H$ gilt dann: $N_G(Q)/Z_G(Q)$ ist eine π-Gruppe.*

Beweis. $N_G(Q)/Z_G(Q)$ ist isomorph zur Gruppe der Automorphismen, die von Elementen aus $N_G(Q)$ in Q induziert werden. Daher genügt es zu zeigen, daß die Ordnungen aller dieser Automorphismen nur Primteiler aus π enthalten. Sei $n \in N_G(Q)$ und q ein beliebiges Element aus Q. Es ist $n = hr$ mit $h \in H$ und $r \in R$; $h^{-1}qh = h_1 \in H$; $n^{-1}qn = r^{-1}h^{-1}qhr = r^{-1}h_1 r \in H$. Da R ausgezeichnet ist, gilt weiter $n^{-1}qn = h_1 r_1^{-1} r$, also $r_1^{-1}r = h_2 \in H$ und $r = r_1 h_2 = h_2 r_2$ mit $r_1, r_2 \in R$. Daraus folgt aber $h_2 = e$ und $r_1 = r$. n und h induzieren also denselben Automorphismus in Q. Seine Ordnung ist ein Teiler von $|H|$ und enthält deshalb nur Primteiler aus π.

Hieraus erhält man den

Satz 1. *Die Sylowgruppe P von G habe in G ein ausgezeichnetes Repräsentantensystem. Dann besitzt P in G ein normales Komplement.*

Beweis. Es sei $|P| = p^\lambda$ und Q eine Untergruppe von G mit p-Potenzordnung. Dann gibt es ein Element $x \in G$ mit $x^{-1}Qx \subseteqq P$. Die Ordnung $|g|$ des Elements $g \in G$ sei nicht durch p teilbar. Liegt g in $N_G(Q)$, so $x^{-1}gx$ in $N_G(x^{-1}Qx)$. Nach dem Lemma ist dann $(x^{-1}gx)^{p^\lambda}$ in $Z_G(x^{-1}Qx)$. Da $(p, |g|) = 1$, liegt auch $x^{-1}gx$ in $Z_G(x^{-1}Qx)$ und daher $g \in Z_G(Q)$. Nach dem Satz von FROBENIUS existiert deshalb ein normales Komplement zu P in G.

3. G heißt *Gruppe mit Sylowturm,* wenn sie eine Untergruppenkette

$$G = G_0 \supset G_1 \supset \ldots \supset G_{r-1} \supset G_r = \{e\}$$

besitzt, in der $G_i \subseteqq | G$ und G_{i-1}/G_i $(i = 1, 2, \ldots, r)$ einer Sylowgruppe von G isomorph ist.

Ist G eine Gruppe mit Sylowturm und U eine Untergruppe von G, so erkennt man bei Betrachtung der Durchschnitte $U_i = G_i \cap U$ $(i = 0, 1, \ldots, r)$, daß auch U eine Gruppe mit Sylowturm ist.

Offenbar besitzt jede nilpotente Gruppe einen Sylowturm. Auch jede überauflösbare Gruppe hat einen Sylowturm, denn zu jeder überauflösbaren Gruppe existiert eine Hauptreihe, in der die Indizes der Größe nach geordnet sind.

Satz 2. *Die Hallgruppe H von G habe in G ein ausgezeichnetes Repräsentantensystem. H sei eine Gruppe mit Sylowturm. Dann besitzt H in G ein normales Komplement.*

Beweis. Ist $|H|$ durch genau n verschiedene Primzahlen teilbar, so enthält H als Sylowturmgruppe sicher eine invariante Hallgruppe H_m, deren Ordnung durch genau $n-m$ $(m = 1, 2, \ldots, n)$ verschiedene Primzahlen teilbar ist. H enthält als auflösbare Gruppe eine Hallgruppe $H^{(m)}$ mit $H = H_m H^{(m)}$, $H_m \cap H^{(m)} = \{e\}$. $H^{(m)}$ hat das ausgezeichnete Repräsentantensystem H_m in H und seine Ordnung ist durch genau m verschiedene Primzahlen teilbar.

Besitzt H in G das ausgezeichnete Repräsentantensystem R und die Untergruppe U von H das ausgezeichnete Repräsentantensystem T in H, so bilden die Elemente des Komplexes TR ein ausgezeichnetes Repräsentantensystem für U in G. Wir betrachten nun die Sylowgruppen von H. Sie sind gleichzeitig Sylowgruppen von G. Diejenigen, die ein ausgezeichnetes Repräsentantensystem in H haben, haben auch eines in G. Nach Satz 1 besitzen sie ein normales Komplement in G.

Jetzt machen wir die Induktionsannahme: Jede Hallgruppe von H, die in H und daher auch in G ein ausgezeichnetes Repräsentantensystem hat und deren

Ordnung höchstens durch m ($1 \leqq m < n$) verschiedene Primzahlen teilbar ist, besitzt in G ein normales Komplement. Behauptung: Dann besitzt auch jede Hallgruppe von H, die in H ein ausgezeichnetes Repräsentantensystem hat und deren Ordnung durch genau $m+1$ verschiedene Primzahlen teilbar ist, ein normales Komplement in G.

Diese Behauptung wird folgendermaßen bewiesen:

Sei U eine Hallgruppe von H mit ausgezeichnetem Repräsentantensystem, deren Ordnung durch genau $m+1$ verschiedene Primzahlen teilbar ist. U ist eine Gruppe mit Sylowturm.

$$U = U_0 \supset U_1 \supset \ldots \supset U_m \supset U_{m+1} = \{e\}$$

sei die entsprechende Untergruppenkette.

U ist als Gruppe mit Sylowturm auflösbar und besitzt daher eine Hallgruppe $U^{(m)}$, für die gilt

$$U = U^{(m)} U_m \quad \text{mit } U^{(m)} \cap U_m = \{e\} \text{ und } U_m \subseteqq | U.$$

U hat ein ausgezeichnetes Repräsentantensystem in H. U_m ist ausgezeichnetes Repräsentantensystem für $U^{(m)}$ in U. Daher hat $U^{(m)}$ ein ausgezeichnetes Repräsentantensystem in H. Ferner ist $U^{(m)}$ Hallgruppe von H. $|U^{(m)}|$ ist genau durch m verschiedene Primzahlen teilbar. Nach der Induktionsannahme besitzt $U^{(m)}$ also ein normales Komplement M_m in G:

$$G = U^{(m)} M_m.$$

Da M_m Normalteiler von G ist, liegt die Sylowgruppe U_m von G in M_m.

π bezeichne die Menge der Primteiler von $|U|$, π' die komplementäre Menge. Ist $|U_m| = p_{m+1}^{\mu}$, so ist p_{m+1} die einzige Primzahl aus π, die $|M_m|$ teilt. Daher liegen alle π-Untergruppen von M_m in U_m oder einer dazu konjugierten Untergruppe. Sei W eine π-Untergruppe von M_m. Dann gibt es ein Element $y \in M_m$ mit $y^{-1} W y = W^* \subseteqq U_m$. Nach dem Lemma ist $N_G(W^*)/Z_G(W^*)$ eine π-Gruppe. Die Ordnungen der Automorphismen von W^*, die durch die Elemente von $N_{M_m}(W^*) \subseteqq N_G(W^*)$ induziert werden, enthalten nur Primteiler aus π. Andererseits muß die Ordnung jedes durch ein Element aus $N_{M_m}(W^*)$ erzeugten Automorphismus von W^* ein Teiler von $|M_m|$ sein. Deshalb ist $N_{M_m}(W^*)/Z_{M_m}(W^*)$ eine p_{m+1}-Gruppe.

Die Ordnung des Elements $Z \in M_m$ sei nicht durch p_{m+1} teilbar. Liegt z in $N_{M_m}(W)$, so $y^{-1} z y$ in $N_{M_m}(y^{-1} W y) = N_{M_m}(W^*)$ und $(y^{-1} z y)^{p_{m+1}^{\mu}}$ in $Z_{M_m}(W^*)$. Da $(p_{m+1}, |Z|) = 1$, ist auch $y^{-1} z y \in Z_{M_m}(W^*)$ und daher $z \in Z_{M_m}(W)$. Nach dem Satz von FROBENIUS existiert deshalb ein normales Komplement M zu U_m in M_m:

$$M_m = U_m M.$$

M ist als Erzeugnis aller Sylowgruppen von M_m, die zu den Primzahlen aus π' gehören, charakteristische Untergruppe von M_m und daher Normalteiler von G. Es ist

$$G = U^{(m)} M_m = U^{(m)} U_m M = UM \quad \text{mit } U \cap M = \{e\}.$$

M ist also normales Komplement zu U in G.

Dieser Induktionsschluß beweist den Satz 2.

Literatur

[1] W. Burnside, *Theory of groups of finite order*, 2. ed. (Cambridge, 1911).
[2] G. Frobenius, Über auflösbare Gruppen. V, *Sitzungsber. preuß. Akad. Wiss.*, **1901**, 1324—1329.
[3] R. Kochendörffer, Ein Satz über Sylowgruppen, *Math. Nachr.*, **17** (1959), 189—194.
[4] G. Zappa, Generalizzazione di un teorema di Kochendörffer, *Le Matematiche*, Catania, **13**, (1958), 61—64.

(Eingegangen am 5. Juni 1964)

Zur Quadratbildung von Polynomen

Von HEINZ SEIBT in Potsdam (DDR)

Prof. L. RÉDEI stellte die Frage, ob es Polynome mit Koeffizienten aus dem Körper Δ der reellen Zahlen gibt, deren Quadrate weniger nichtverschwindende Glieder haben als das jeweilige Ausgangspolynom (vgl. [1] und [2]). Man bezeichne mit $f_N(x) \in \Delta[x]$ ein Polynom mit N von 0 verschiedenen Gliedern und mit $Q(f_N(x))$ die Anzahl der nichtverschwindenden Glieder von $f_N^2(x)$. Weiterhin sei

$$Q(N) = \min Q(f_N(x)),$$

worin $f_N(x)$ alle Polynome aus $\Delta[x]$ mit N von 0 verschiedenen Gliedern durchläuft. Das oben genannte Problem entspricht dann der Frage, ob es natürliche Zahlen N gibt, für die

$$Q(N) < N \quad \text{bzw.} \quad q(N) = \frac{Q(N)}{N} < 1$$

gilt.

A. RÉNYI untersucht in [2] weitere Eigenschaften der Funktionen $Q(N)$ bzw. $q(N)$ und zeigt u. a.

$$q(4n+1) \leqq \frac{28}{29} < 1 \quad \text{für} \quad n \geqq 7.$$

Dazu konstruiert er für alle natürlichen Zahlen $N = 4n+1$ Polynome $A_{4n+1}(x)$ mit N von 0 verschiedenen Gliedern, deren Quadrate für $n \geqq 7$ aus weniger als N von 0 verschiedenen Gliedern bestehen. Dabei ist $N = 29$ die kleinste nach [2] bekannte Zahl mit $Q(N) < N$ (vgl. aber Fußnote [1])).

W. VERDENIUS zeigt in [3] durch Konstruktion gewisser Polynome die Abschätzungen

$$Q(N) < \frac{1}{7}(312 \cdot (N-1)^{^{13}\log 8} - 18) \quad \text{bzw.} \quad Q(N) < \frac{1}{5}(162 \cdot (N-1)^{^{9}\log 6} - 12).$$

Daraus folgt $Q(N) < N$ angenähert für $N > 4{,}8 \cdot 10^8$ bzw. für $N > 2{,}3 \cdot 10^8$.

Ziel unserer Ausführungen ist es, den Bereich derjenigen natürlichen Zahlen N, für die $Q(N) < N$ gilt, näher anzugeben. Dazu sprechen wir den folgenden Satz aus.

Satz. *Mit Ausnahme der (nicht geklärten) Zahlen* $N = 26, 27, 34$ *gilt*

$$Q(N) < N$$

für alle natürlichen Zahlen $N \geqq 21.$

Wir werden den Beweis so führen, daß wir zunächst gewisse Serien von Polynomen $f_N(x)$ mit $Q(f_N(x)) < N$ konstruieren. Diese werden die im Satz genannten Werte für N bereits bis auf wenige Ausnahmen liefern, welche dann durch Einzelbetrachtungen erledigt werden können.*)

1.

Grundlage unserer Untersuchungen sind die folgenden von Rényi in [2] eingeführten Polynome $S_n(x)$ und $P_{2n+1}(x)$: Für jede natürliche Zahl $n \geqq 2$ bezeichne $S_n(x)$ die ersten n Glieder der Entwicklung

$$\sqrt{1+2x} = 1 + x + a_2 x^2 + \dots + a_{n-1} x^{n-1} + \dots,$$

also

$$S_n(x) = \sum_{k=0}^{n-1} a_k x^k \quad \text{mit} \quad a_k = \binom{\frac{1}{2}}{k} \cdot 2^k \qquad (k = 0, 1, \dots, n-1).$$

Dann enthält $S_n^2(x)$ genau $n+1$ von 0 verschiedene Glieder, nämlich die mit den Exponenten $0, 1, n, n+1, \dots, 2n-2$. Weiter sei für jede natürliche Zahl $n \geqq 2$

$$\lambda = \sqrt{\frac{a_{n-1}}{a_{n+1}}},$$

und $P_{2n+1}(x)$ sei das aus $2n+1$ Gliedern bestehende Polynom

$$P_{2n+1}(x) = \sum_{j=0}^{n-1} b_j (x^j + x^{2n-j}) + b_n x^n$$

mit

$$b_j = \binom{\frac{1}{2}}{j} \cdot \lambda^j 2^j \qquad (j = 0, 1, \dots, n).$$

Dabei sind die b_j gerade die Koeffizienten der ersten $n+1$ Glieder der Entwicklung

$$\sqrt{1+2\lambda x} = 1 + b_1 x + \dots + b_n x^n + \dots.$$

Das Quadrat $P_{2n+1}^2(x)$ hat genau $2n+1$ von 0 verschiedene Glieder, und zwar diejenigen mit den Exponenten $0, 1, n+2, \dots, 3n-2, 4n-1, 4n$.

2.

In Verallgemeinerung eines Konstruktionsverfahrens von Rényi bilden wir folgende Polynome

$$f(x) = P_{2m+1}(x) \cdot S_n(\mu x^{2m}) = 1 + c_1 x + c_2 x^2 + \dots + c_{2mn} x^{2mn},$$

*) Für freundliche Unterstützung beim Entstehen dieser Arbeit möchte ich Herrn Prof. Dr. H. Weinert meinen Dank aussprechen.

die von den natürlichen Zahlen $m \geqq 2$ und $n \geqq 2$ sowie von einem reellen Parameter $\mu \neq 0$ abhängen [1]) und zeigen:

Hilfssatz 1. A) *Das Polynom $f(x)$ hat genau $2mn+1$ von 0 verschiedene Glieder, wenn $\mu \neq \frac{k}{2k-3}$ $(k=1, 2, \ldots, n-1)$ gewählt wird; dagegen reduziert sich die Gliederzahl auf $2mn$, wenn μ einen dieser Werte $\frac{k}{2k-3}$ annimmt.*

B) *Unabhängig von der Wahl von μ hat $f^2(x)$ für $m=2$ höchstens $3n+7$, für $m \geqq 3$ höchstens $2mn+2m-3n+5$ von 0 verschiedene Glieder.*

Beweis. A) In $f(x)=P_{2m+1}(x) \cdot S_n(\mu x^{2m})$ erhalten wir die Glieder 1, $c_{2mn}x^{2mn}$ und $c_i x^i$ mit $i \not\equiv 0\ (2m)$ genau einmal, deren Koeffizienten verschwinden also nicht. Für die Koeffizienten der Glieder $c_{2mk}x^{2mk}$ $(k=1, 2, \ldots, n-1)$ gilt

$$c_{2mk} = \left[\binom{\frac{1}{2}}{k} \cdot 2^k \mu^k + \binom{\frac{1}{2}}{k-1} \cdot 2^{k-1} \mu^{k-1}\right] =$$

$$= \left(\frac{\frac{1}{2}-(k-1)}{k} 2\mu + 1\right) \cdot \binom{\frac{1}{2}}{k-1} \cdot 2^{k-1} \mu^{k-1} = \frac{(3-2k)\mu + k}{k} \binom{\frac{1}{2}}{k-1} \cdot 2^{k-1} \mu^{k-1}.$$

Offensichtlich verschwindet c_{2mk} genau dann, wenn $(3-2k)\mu + k = 0$ gilt, d. h. für

$$\mu = \frac{k}{2k-3} \qquad (k=1, 2, \ldots, n-1).$$

Wir können daraus ersehen, daß jedes k einen Wert des Parameters μ eindeutig bestimmt und daß verschiedene Werte von k auch verschiedene Werte von μ bedingen. Nur für gewisse Werte von μ aus dem Intervall $[-1, 2]$ verschwindet demnach genau ein Koeffizient c_{2mk}.

Damit hat $f(x)$ je nach der Wahl von μ: genau $2mn+1$ Glieder, falls $\mu \neq \frac{k}{2k-3}$, bzw. genau $2mn$ Glieder, falls für μ einer der Werte $\frac{k}{2k-3}$ gewählt wird.

B) Wie wir bereits erwähnt haben, besteht $P^2_{2m+1}(x)$ nur aus Gliedern mit den Exponenten $0, 1, m+2, \ldots, 2m, 1+2m, \ldots, m-2+2m, 2m-1+2m, 2m+2m$; in $S^2_n(\mu x^{2m})$ kommen höchstens Glieder mit den Exponenten $2mk$ $(k=0, 1, n, \ldots, 2n-2)$ vor. Wir wollen nun die Anzahl der Glieder von $f^2(x)$ bestimmen. Dazu multiplizieren wir der Reihe nach die Glieder von $P^2_{2m+1}(x)$ mit allen Gliedern von $S^2_n(\mu x^{2m})$. Die sich dabei ergebenden Exponenten ordnen wir in einem Schema an, wobei wir die Numerierung der einzelnen Zeilen den Exponenten der Glieder von $P^2_{2m+1}(x)$ entsprechend wählen.

---

[1]) Rényi betrachtet nur die Polynome mit $m=2$ und (nach einer brieflichen Mitteilung) $\mu=3$, die gemäß dem folgenden Hilfssatz 1 aus $4n+1$ von 0 verschiedenen Gliedern bestehen und deren Quadrate für $n \geqq 7$ weniger als $4n+1$ Glieder besitzen. Durch einen Druckfehler stehen in seiner Arbeit jedoch die Polynome mit $m=2$ und $\mu=1$, die nur $4n$ von 0 verschiedene Glieder haben und deren Quadrate erst für $n \geqq 8$ weniger als $4n$ Glieder aufweisen.

$$0) \qquad 2mk \quad (k=0,1,n,\ldots,2n-2)$$

also $n+1$ Glieder;

$$1) \qquad 1+2mk \quad (k=0,1,n,\ldots,2n-2)$$

weitere $n+1$ Glieder;

$$\begin{array}{ll} m+2) & m+2+2mk \quad (k=0,1,n,\ldots,2n-2) \\ m+3) & m+3+2mk \quad (k=0,1,n,\ldots,2n-2) \\ \vdots & \quad\vdots \\ 2m-1) & 2m-1+2mk \quad (k=0,1,n,\ldots,2n-2) \end{array}$$

das sind $m-2$ Zeilen mit je $n+1$ Gliedern, also $(m-2)(n+1)$ Glieder;

$$2m) \qquad 2mk \quad (k=1,2,n+1,\ldots,2n-1)$$

außer $k=2$ und $k=2n-1$ sind diese Glieder bereits in der Zeile 0) erfaßt, also 2 weitere Glieder;

$$\begin{array}{ll} 1+2m) & 1+2mk \quad (k=1,2,n+1,\ldots,2n-1) \\ & \quad\vdots \\ m-2+2m) & m-2+2mk \quad (k=1,2,n+1,\ldots,2n-1), \end{array}$$

das sind $m-2$ Zeilen zu je $n+1$ Gliedern, von denen für $m \geqq 3$ bereits $n-1$ Glieder der Zeile $1+2m$) schon in der Zeile 1) gezählt wurden, also $m=2$: 0 Glieder, $m \geqq 3$: $(m-2)(n+1)-(n-1)$ Glieder;

$$2m-1+2m) \qquad 2m-1+2mk \quad (k=1,2,n+1,\ldots,2n-1)$$

von diesen $n+1$ Gliedern sind für $m \geqq 3$ alle außer $k=2$ und $k=2n-1$ bereits in Zeile $2m-1$) gezählt worden, also $m=2$: $n+1$ Glieder, $m \geqq 3$: 2 Glieder;

$$2m+2m) \qquad 2mk \quad (k=2,3,n+2,\ldots,2n)$$

das sind weitere 2 Glieder, da alle außer $k=3$ und $k=2n$ schon in Zeile 0) bzw. $2m$) gezählt wurden.

Durch Addition der Gliederzahlen in den einzelnen Zeilen erhalten wir:

$m=2$: $f^2(x)$ hat höchstens $3n+7$ Glieder,
$m \geqq 3$: $f^2(x)$ hat höchstens $2mn+2m-3n+5$ Glieder.

Aus diesem Hilfssatz 1 ergeben sich die folgenden Abschätzungen, die für $n \geqq 2$ gelten:

$m=2$: $Q(4n+1) \leqq 3n+7$ bzw. $Q(4n) \leqq 3n+7$,
$m \geqq 3$: $Q(2mn+1) \leqq 2mn+2m-3n+5$ bzw. $Q(2mn) \leqq 2mn+2m-3n+5$.

Insbesondere gilt dann

$m=2$: $Q(4n+1) < 4n+1$ für $n \geqq 7$ bzw. $Q(4n) < 4n$ für $n \geqq 8$,

$m \geqq 3$: $Q(2mn+1) < 2mn+1$ für $n > \frac{2m+4}{3}$ bzw. $Q(2mn) < 2mn$ für $n > \frac{2m+5}{3}$.

Dabei lehren die beiden ersten Zeilen, daß für alle Zahlen $N\equiv 0$ (4) bzw. $N\equiv 1$ (4) ab $N=32$ bzw. ab $N=29$ die Beziehung $Q(N)<N$ gilt[2]). Darüber hinaus interessieren uns auf Grund des Folgenden aus den Serien mit $m=3$ nur die Werte $N=24=2\cdot 3\cdot 4$, $N=25=2\cdot 3\cdot 4+1$, $N=30=2\cdot 3\cdot 5$, $N=31=2\cdot 3\cdot 5+1$, $N=42=$ $=2\cdot 3\cdot 7$, $N=43=2\cdot 3\cdot 7+1$, für die ebenfalls die Beziehung $Q(N)<N$ zutrifft.

3.

Um uns einen entsprechenden Überblick über die Zahlen $N\equiv 2$ (4) und $N\equiv 3$ (4) zu verschaffen, betrachten wir die Polynome

$$f(x) = P_5(x^3)\cdot S_n(vx^4) = 1+d_1x+d_2x^2+\ldots+d_{4(n+2)}x^{4(n+2)},$$

die von der natürlichen Zahl $n\geqq 2$ und einem reellen Parameter $v\neq 0$ abhängen. Dabei verschwinden von vornherein einige der Koeffizienten (etwa d_1, d_2). Wir zeigen

Hilfssatz 2. A) *Das Polynom $f(x)$ hat genau $4n+3$ von 0 verschiedene Glieder, wenn man*

$$v\neq \sqrt[3]{\frac{k(k-1)(k-2)}{(2k-7)(2k-5)(2k-3)}} \qquad (k=3, 4, \ldots, n-1)$$

wählt, hingegen verringert sich die Anzahl der Glieder auf $4n+2$, wenn v einen dieser Werte annimmt.

B) *Unabhängig von der Wahl von v hat $f^2(x)$ höchstens $3n+13$ von 0 verschiedene Glieder.*

Beweis. A) Das Polynom $P_5(x^3)$ hat nur Glieder mit den Exponenten $0, 3, 6\equiv 2(4)$, $9\equiv 1(4)$ und $12\equiv 0(4)$; das Polynom $S_n(vx^4)$ hat nur Glieder mit den Exponenten $4k$ $(k=0, 1, \ldots, n-1)$. Für $f(x)$ erhalten wir damit zunächst je n Glieder d_ix^i mit $i\equiv 3(4)$, bzw. $i\equiv 2(4)$, bzw. $i\equiv 1(4)$. Hinzu kommen die Glieder $c_{4k}x^{4k}$ mit $k=0, 1, 2, n, n+1, n+2$. Das sind insgesamt $3n+6$ Glieder mit nichtverschwindenden Koeffizienten. Für die Koeffizienten der restlichen $n-3$ Glieder $d_{4k}x^{4k}$ mit $k=3, 4, \ldots, n-1$ gilt

$$d_{4k} = \left[\binom{\frac{1}{2}}{k-3} 2^{k-3}v^{k-3}+\binom{\frac{1}{2}}{k}2^kv^k\right] =$$

$$= \left[1+\frac{(\frac{1}{2}-(k-3))(\frac{1}{2}-(k-2))(\frac{1}{2}-(k-1))}{(k-2)(k-1)k}\cdot 2^3v^3\right]\binom{\frac{1}{2}}{k-3}2^{k-3}v^{k-3} =$$

$$= \frac{(k-2)(k-1)k+(7-2k)(5-2k)(3-2k)v^3}{k(k-1)(k-2)}\binom{\frac{1}{2}}{k-3}2^{k-3}v^{k-3}.$$

---

[2]) Das Ergebnis $Q(4n+1)\leqq 3n+7<4n+1$ für $n\geqq 7$ ist gemäß der vorstehenden Fußnote bereits bei Rényi in [2] zu finden.

Demnach verschwindet d_{4k} genau dann, wenn

$$k(k-1)(k-2)+(7-2k)(5-2k)(3-2k)v^3 = 0$$

gilt, d. h. für

$$v = \sqrt[3]{\frac{k(k-1)(k-2)}{(2k-3)(2k-5)(2k-7)}} \qquad (k=3, 4, \ldots, n-1).$$

Diese Formel lehrt, daß jeder der in Frage kommenden Werte von k einen Wert des Parameters v bestimmt und daß zu verschiedenen k auch verschiedene v gehören. Es verschwindet also nur für gewisse Werte v aus dem Intervall $\left[-\sqrt[3]{2}, \frac{2}{\sqrt[3]{5}}\right]$ genau ein Koeffizient d_{4k}. Somit hat das Polynom $f(x)$ entsprechend der Wahl von v genau $4n+3$ bzw. genau $4n+2$ von 0 verschiedene Glieder.

B) Nach dem Vorigen besitzt $P_5^2(x^3)$ nur Glieder mit den Exponenten 0, 3, $12=4\cdot 3$, $3\cdot 7=21=1+4\cdot 5$, $3\cdot 8=24=4\cdot 6$; dagegen hat $S_n^2(vx^4)$ höchstens Glieder mit Exponenten $4k$ $(k=0, 1, n, \ldots, 2n-2)$. Entsprechend Beweisteil B) von Hilfssatz 1 verschaffen wir uns in einem Schema einen Überblick über die Glieder von $f^2(x)$:

0) $\qquad 4k \qquad (k=0, 1, n, \ldots, 2n-2)$,

das sind $n+1$ Glieder;

3) $\qquad 3+4k \qquad (k=0, 1, n, \ldots, 2n-2)$,

also weitere $n+1$ Glieder;

4·3) $\qquad 4k \qquad (k=3, 4, n+3, \ldots, 2n+1)$,

das sind höchstens 5 weitere Glieder, nämlich die für $k=3, 4, 2n-1, 2n, 2n+1$;

1+4·5) $\qquad 1+4k \qquad (k=5, 6, n+5, \ldots, 2n+3)$,

also $n+1$ weitere Glieder;

4·6) $\qquad 4k \qquad (k=6, 7, n+6, \ldots, 2n+4)$,

das sind höchstens 5 weitere Glieder, und zwar die mit $k=6, 7, 2n+2, 2n+3, 2n+4$.

Durch Addition der jeweiligen Gliederzahlen ergibt sich die Behauptung: $f^2(x)$ hat höchstens $3n+13$ Glieder.

Damit erhalten wir aus Hilfssatz 2 die folgenden Abschätzungen für $Q(N)$, die für $n \geqq 2$ gelten:

$$Q(4n+3) \leqq 3n+13, \quad Q(4n+2) \leqq 3n+13.$$

Daraus folgt insbesondere

$$Q(4n+3) < 4n+3 \quad \text{für } n \geqq 11,$$

$$Q(4n+2) < 4n+2 \quad \text{für } n \geqq 12.$$

Die beiden letzten Zeilen bestätigen die Richtigkeit der Beziehung $Q(N) < N$ somit für alle Zahlen $N \equiv 2(4)$ bzw. $N \equiv 3(4)$ ab $N=50$ bzw. ab $N=47$.

4.

Um den Beweis unseres Satzes abzuschließen, geben wir nun noch für die durch die beiden Hilfssätze nicht erfaßten Zahlen N geeignete Polynome an, die auch für diese Zahlen die Beziehung $Q(N) < N$ bestätigen:

$$
\begin{array}{lll}
N = 46: & f_{46}(x) = P_5(x^3)\cdot P_{11}\left(\sqrt[6]{\dfrac{125}{168}}\,x^4\right) & \text{mit} \quad Q(f_{46}(x)) \leqq 45, \\
N = 39: & f_{39}(x) = P_5(x^3)\cdot P_9\,(x^4) & \text{mit} \quad Q(f_{39}(x)) \leqq 37, \\
N = 38: & f_{38}(x) = P_5(x^3)\cdot P_9\left(\sqrt[6]{\dfrac{4}{7}}\,x^4\right) & \text{mit} \quad Q(f_{38}(x)) \leqq 37, \\
N = 35: & f_{35}(x) = P_5(x)\cdot P_7\,(x^7) & \text{mit} \quad Q(f_{35}(x)) \leqq 27, \\
N = 28: & f_{28}(x) = P_5(x)\cdot P_7\left(-\dfrac{2}{\sqrt{5}}\,x^4\right) & \text{mit} \quad Q(f_{28}(x)) \leqq 27, \\
N = 23: & f_{23}(x) = P_5(x^3)\cdot P_5\,(x^4) & \text{mit} \quad Q(f_{23}(x)) \leqq 21, \\
N = 22: & f_{22}(x) = P_5(x^3)\cdot P_5\,(-\sqrt[6]{2}\,x^4) & \text{mit} \quad Q(f_{22}(x)) \leqq 21, \\
N = 21: & f_{21}(x) = P_5(x)\cdot P_5\left(\dfrac{7}{4}\sqrt{2}x\right) & \text{mit} \quad Q(f_{21}(x)) \leqq 20.
\end{array}
$$

Die in dieser Tabelle enthaltenen Behauptungen über N und $Q(f_N(x))$ lassen sich mit ähnlichen Gedankengängen wie in den Beweisen zu unseren Hilfssätzen zeigen; wir verzichten jedoch darauf, dieses auszuführen. Übrigens lassen sich auch zu diesen Polynomen gewisse Serien aufstellen. Diese Serien sind aber auf Grund der mit ihnen erfaßten Zahlen N zu einem Beweis unseres Satzes weniger geeignet, dagegen liefern sie im allgemeinen bessere Abschätzungen für $Q(N)$ bzw. $q(N)$. Beispielsweise gewinnt man aus $P_5(x)\cdot P_9(x^7)$ die Abschätzung $q(45) \leqq \frac{33}{45}$, während das Polynom $P_5(x)\cdot S_{11}(3x^4)$ aus Hilfssatz 1 nur $q(45) \leqq \frac{40}{45}$ ergibt.

Literatur

[1] L. Rédei, *Algebra*. I (Leipzig, 1959), S. 271.

[2] A. Rényi, On the minimal number of terms of the square of a polynomial, *Acta Math. Hung.*, **1** (1947), 30–34.

[3] W. Verdenius, On the number of terms of the square and the cube of polynomials, *Indag. Math.*, **11** (1949), 459–465.

[4] P. Erdős, On the number of terms of the square of a polynomial, *Nieuw Arch. Wisk.*, **23** (1949), 63–65.

(Eingegangen am 21. Dezember 1963)

Zur Theorie der Algebren und monomialen Ringe

Von HANNS JOACHIM WEINERT in Potsdam (DDR)

Einleitung

Es sei R ein Ring, dessen Modul R^+ Linksvektorraum über einem Ring $\mathfrak{R}$ ist, so daß also für beliebige a, b aus $\mathfrak{R}$ und α, β aus R

$$(1) \qquad a(\alpha+\beta) = a\alpha+a\beta, \; (a+b)\alpha = a\alpha+b\alpha, \; (ab)\alpha = a(b\alpha)$$

gilt und eine (linear unabhängige) Basis $\{\omega_i\}$ beliebiger Mächtigkeit existiert; letzteres besagt, daß sich jedes Element $\alpha \in R$ eindeutig gemäß

$$\alpha = \sum_i a_i \omega_i, \quad \text{fast alle} \quad a_i = 0$$

linear kombinieren läßt. Üblicherweise nennt man R eine (linksseitige) Algebra über $\mathfrak{R}$, wenn darüber hinaus stets

$$(2) \qquad a(\alpha\beta) = (a\alpha)\beta = \alpha(a\beta)$$

erfüllt ist. Dies hat den Vorteil, daß dann bezüglich jeder Basis $\{\omega_i\}$ von R^+ die Multiplikation in R bereits durch die Produkte

$$(3) \qquad \omega_i \omega_j = \sum_k c_{ij}^{(k)} \omega_k$$

bzw. durch die Strukturkonstanten $c_{ij}^{(k)}$ bestimmt ist. Doch ergibt sich aus (2) im allgemeinen die Notwendigkeit, mit kommutativen Grundringen $\mathfrak{R}$ zu arbeiten, wodurch wichtige Klassen algebraischer Strukturen nicht erfaßt werden. Aus diesem Grunde führte PICKERT in [4] den Begriff der *Algebra im weiteren Sinne* (i. w. S.) ein, wobei (2) durch eine schwächere Forderung ersetzt wird. Einen anderen Weg geht RÉDEI in [5], indem er R als Links- und Rechtsvektorraum über $\mathfrak{R}$ betrachtet und den Begriff der *Doppelalgebra* einführt, ohne jedoch eine entsprechende Verallgemeinerung des Begriffes der (linksseitigen) Algebra vorzunehmen. Schließlich lassen sich aber auch auf diese Weise eine Reihe von Strukturen noch nicht erfassen, für die eine einheitliche Behandlung zusammen mit den bisher genannten Algebrenbegriffen wünschenswert wäre; als Beispiele hierzu nennen wir die „nichtkommutativen" Polynomringe von ORE (vgl. [3]) und die von RÉDEI in [5], § 65 bzw. von BÓDI in [2] eingeführten verschränkten Produkte, auf die wir auch im Text noch einmal zu sprechen kommen.

In der vorliegenden Arbeit stellen wir daher in § 1 zunächst einige allgemeine Aussagen über Ringe R zusammen, die Vektorräume über $\mathfrak{R}$ sind, welche die entsprechenden Aussagen für die oben genannten Algebrenbegriffe enthalten. Weiterhin werden wir durch diese Betrachtungen eigentlich zwangsläufig auf den PICKERTschen Begriff der Algebra i. w. S. geführt, der einen glücklichen Kompromiß zwischen Einfachheit und Allgemeinheit darstellt (vgl. § 2). Bei der Behandlung dieser Algebren können wir uns freilich im Hinblick auf [4] kurz fassen. In § 3 zeigen wir, daß der RÉDEIsche Begriff der Doppelalgebra mit dem der Algebra i. w. S. in folgendem Sinne gleichwertig ist: Jede Doppelalgebra ist stets auch Algebra i. w. S., während aus jeder Algebra i. w. S. durch geeignete Definition einer Rechtsoperatoranwendung der Elemente von $\mathfrak{R}$ eine Doppelalgebra entsteht, falls $\mathfrak{R}$ ein Einselement enthält.

Schließlich entwickeln wir in § 4 die Grundlagen einer allgemeinen Theorie der *monomialen Ringe*. Dabei stützen wir unsere Begriffsbildungen zunächst nur auf § 1, geben ein Assoziativitätskriterium und führen allgemeine Halbgruppenringe als spezielle monomiale Ringe ein. Es stellt sich aber heraus, daß jeder assoziative monomiale Ring entweder selbst ein allgemeiner Halbgruppenring ist oder in einfacher Weise als Restklassenring eines solchen Halbgruppenringes gewonnen werden kann. Diese Aussagen gelten (natürlich mit gewissen Vereinfachungen) erst recht, wenn die betrachteten monomialen Ringe sogar Algebren i. w. S. sind, womit wir im wesentlichen zu dem in [5], § 66 durchgeführten Fall gelangen.

In einer Fortsetzung dieser Arbeit werden wir auf das von RÉDEI a. a. O. gestellte und zum Teil behandelte Problem eingehen, einen Überblick über alle monomialen Ringe mit einem bestimmten Faktorensystem zu gewinnen.

§ 1. Vektorraumringe

Der Kürze halber wollen wir einen Ring R, der zugleich Linksvektorraum über einem Ring $\mathfrak{R}$ ist (vgl. Einleitung), einen Vektorraumring über $\mathfrak{R}$ nennen. Dabei sei der Grundring $\mathfrak{R}$ stets assoziativ, während wir für R auch nichtassoziative Ringe zulassen; weiterhin enthalte $\mathfrak{R}$ stets wenigstens ein Rechtseinselement e_r.[1]). Ist $\{\omega_i\}$ eine beliebige Basis von R über $\mathfrak{R}$, wobei i eine Indexmenge I durchläuft, so gilt für das Produkt zweier „Monome" $a\omega_i$ und $b\omega_j$ aus R

(4) $$a\omega_i \cdot b\omega_j = \sum_{k \in I} F(a, b, i, j, k)\,\omega_k,$$

wobei die Koeffizienten $F(a, b, i, j, k) \in \mathfrak{R}$ von allen angegebenen Argumenten abhängen können und natürlich für feste Wahl von a, b, i und j nur jeweils endlich viele verschieden von 0 sind. Weiterhin kann dann das Produkt beliebiger Elemente aus R gemäß

(5) $$\left(\sum a_i\omega_i\right)\left(\sum_j b_j\omega_j\right) = \sum_{i,j,k} F(a_i, b_j, i, j, k)\,\omega_k$$

---

[1]) Bekanntlich ist die Existenz eines Rechtseinselementes in $\mathfrak{R}$ notwendig und hinreichend für die Existenz von (Links-) Vektorräumen über $\mathfrak{R}$, vgl. etwa [4].

angegeben werden. Aus dem distributiven Gesetz folgen noch die Regeln

(6) $$\begin{aligned} F(a+a', b, i, j, k) &= F(a, b, i, j, k) + F(a', b, i, j, k) \\ F(a, b+b', i, j, k) &= F(a, b, i, j, k) + F(a, b', i, j, k), \end{aligned}$$

woraus sich insbesondere ergibt:

(7) $$F(0, b, i, j, k) = F(a, 0, i, j, k) = 0.$$

Allgemein wollen wir eine Funktion $F(a, b, i, j, k)$, welche jedem Argumentsystem $a \in \mathfrak{R}, b \in \mathfrak{R}, i \in I, j \in I, k \in I$ eindeutig ein Element aus $\mathfrak{R}$ zuordnet, eine *Strukturfunktion* (bezüglich $\mathfrak{R}$ und I) nennen, wenn sie der im Anschluß an (4) formulierten Endlichkeitsbedingung und den Regeln (6) genügt.

Satz 1. *Ist R ein Vektorraumring über $\mathfrak{R}$ mit der Basis $\{\omega_\iota\}_{\iota \in I}$, so wird durch das Produkt der Monome gemäß* (4) *eine Strukturfunktion gegeben, welche die Multiplikation in R nach* (5) *festlegt. Der Vektorraumring R ist also durch $\mathfrak{R}$, die Mächtigkeit von I und seine Strukturfunktion $F(a, b, i, j, k)$ bis auf Isomorphie eindeutig bestimmt. Umgekehrt entsteht zu jedem Ring $\mathfrak{R}$, einer Indexmenge I und einer Strukturfunktion bezüglich $\mathfrak{R}$ und I ein Vektorraumring R mit dieser Strukturfunktion, indem man in einen $\mathfrak{R}$-Vektorraum R^+ mit einer Basis $\{\omega_\iota\}_{\iota \in I}$ eine Multiplikation durch* (5) *einführt.*

Nach dem Vorangegangenen ist nur noch die letzte Behauptung zu beweisen. Dabei ist klar, daß durch (5) je zwei Elementen von R^+ eindeutig ein Element aus R^+ als Produkt zugeordnet wird, und aus (5) auch (4) folgt. Die Gültigkeit des distributiven Gesetzes ergibt sich unmittelbar auf Grund der für die Strukturfunktion vorausgesetzten Regeln (6).

Satz 2. *Ein Vektorraumring R über $\mathfrak{R}$ ist genau dann assoziativ, wenn seine Strukturfunktion den Assoziativitätsbedingungen*

(8) $$\sum_t F\big(F(a, b, i, j, t), c, t, k, l\big) = \sum_t F\big(a, F(b, c, j, k, t), i, t, l\big)$$

für alle a, b und c aus $\mathfrak{R}$ und alle i, j, k und l aus I genügt.

Beweis. Unter Verwendung von (4) und (5) erhält man, daß

$$(a\omega_i b\omega_j)\, c\omega_k = a\omega_i(b\omega_j\, c\omega_k)$$

für alle a, b, c, i, j und k gerade mit der Bedingung (8) gleichwertig ist. Daraus folgt aber unter Verwendung des distributiven Gesetzes bzw. nach (6) bereits die Assoziativität der Multiplikation in R.

Als erstes Beispiel für diese allgemeinen Überlegungen betrachten wir Oresche Polynomringe (vgl. [3]). Dazu sei $\mathfrak{R}$ ein Ring mit Einselement[2]), η ein Endomorphismus von $\mathfrak{R}$ und δ eine zu diesem Endomorphismus korrespondierende Ableitung in $\mathfrak{R}$, d. h. eine eindeutige Abbildung von $\mathfrak{R}$ in sich, die

$$(a+b)^\delta = a^\delta + b^\delta \quad \text{und} \quad (ab)^\delta = a^\eta b^\delta + a^\delta b$$

---

[2]) Ore betrachtet als Grundbereiche nur Körper, doch macht diese Verallgemeinerung unsere Überlegungen nicht schwieriger.

für alle a und b aus $\mathfrak{R}$ erfüllt. ORE betrachtet dann den Ring R aller Polynome $\Sigma\, a_i x^i$ mit $a_i \in \mathfrak{R}$, dessen Multiplikation sich von der üblichen im wesentlichen durch $xb = b^\eta x + b^\delta$ unterscheidet. Daraus folgt allgemeiner

$$ax^i bx^j = aS_{i,0}(b)\, x^{i+j} + aS_{i,1}(b)\, x^{i+j-1} + \ldots + aS_{i,i}(b)\, x^j$$

mit

$$S_{i,t}(b) = \sum b^{\eta^{i-t}\delta^t},$$

wobei diese Summen über alle $\binom{i}{t}$ Permutationen (mit Wiederholung) der $i-t$ Exponenten η und der t Exponenten δ zu erstrecken ist. Allerdings wird in [3] die Frage, ob es solche Ringe überhaupt gibt, d. h. ob zu einem Ring $\mathfrak{R}$ mit einem Paar (η, δ) derartiger Abbildungen ein (assoziativer) Ring R mit der obigen Multiplikation existiert, gar nicht gestellt. Wir beantworten diese Frage positiv, indem wir für $\mathfrak{R}$ mit der Indexmenge $I = \{0, 1, 2, \ldots\}$ eine Strukturfunktion gemäß

$$F(a, b, i, j, k) = aS_{i,\, i+j-k}(b) \quad \text{mit} \quad S_{i,t}(b) = 0 \quad \text{für} \quad t < 0 \text{ bzw. } t > i$$

einführen, welche die Endlichkeitsbedingung und (6) ersichtlich erfüllt. Nach Satz 1 gibt es dann einen entsprechenden Vektorraumring R über $\mathfrak{R}$ mit der Basis $\{\omega, \omega_1, \ldots\}$, der bis auf die Bezeichnung ω_i für x^i gerade der OREsche Polynomring ist. Dieser Ring R ist nach Satz 2 stets assoziativ, denn es gilt für alle a, b, c aus $\mathfrak{R}$ und alle i, j, k, l aus I:

$$\sum_t F\big(F(a, b, i, j, t), c, t, k, l\big) = \sum_t aS_{i,i+j-t}(b)\, S_{t,t+k-l}(c) =$$

$$= \sum_s aS_{i,i+s-l}\big(bS_{j,j+k-s}(c)\big) = \sum_s F\big(a, F(b, c, j, k, s), i, s, l\big).$$

Wir bemerken hierzu lediglich, daß nur für $k \leqq l \leqq i+j+k$ von 0 verschiedene Summanden auftreten, und die Summationen aus dem gleichen Grunde auf

$$\max(j, l-k) \leqq t \leqq i+j$$
$$\max(k, l-i) \leqq s \leqq \min(l, j+k)$$

beschränkt werden können. Im übrigen erfordert der Nachweis dieser Gleichheit ein ausführliches Ausschreiben beider Seiten, worauf wir hier jedoch nicht eingehen wollen.

Erheblich einfacher werden diese Überlegungen für den Spezialfall der OREschen Polynomringe[3]), wo für η ein nichttrivialer Endomorphismus von $\mathfrak{R}$, für δ jedoch der Nullendomorphismus genommen wird. Dann erhalten wir

$$F(a, b, i, j, k) = \begin{cases} ab^{\eta^i} & \text{für} \quad k = i+j \\ 0 & \text{für} \quad k \neq i+j \end{cases}$$

und damit

$$a\omega_i b\omega_j = ab^{\eta^i}\omega_{i+j}, \quad \text{also} \quad ax^i bx^j = ab^{\eta^i} x^{i+j},$$

[3]) Solche Ringe werden z. B. in [1] und [6] herangezogen. Übrigens ist auch der Fall, wo η die identische Abbildung und δ eine Ableitung im üblichen Sinne ist, entsprechend leicht zu behandeln.

und die obige Assoziativitätsbedingung reduziert sich auf

$$ab^{\eta^i}c^{\eta^{i+j}} = a(bc^{\eta^j})^{\eta^i}.$$

Als zweites Beispiel behandeln wir eine Klasse von Vektorraumringen, welche die verschränkten Produkte im Sinne von [2] bzw. [5] als Spezialfälle enthält. Hierzu sei $\mathfrak{R}$ ein Ring und $H=\{\alpha, \beta, \ldots\}$ eine Halbgruppe. Jedem $\alpha \in H$ wird ein Endomorphismus von $\mathfrak{R}$ zugeordnet, wofür wir kurz $a \to \alpha a$ mit

$$\alpha(a+b) = \alpha a + \alpha b, \quad \alpha(ab) = \alpha a\ \alpha b$$

schreiben. Weiter sei ein „Faktorensystem" $c_{\alpha,\beta} \in \mathfrak{R}$ gegeben, welches zusammen mit dem „Endomorphismensystem" αa den Relationen

$$c_{\alpha,\beta}\ c_{\alpha\beta,\gamma} = \alpha(c_{\beta,\gamma})\ c_{\alpha,\beta\gamma}$$

$$c_{\alpha,\beta}\ (\alpha\beta)c = \alpha(\beta c)\ c_{\alpha,\beta}$$

für alle $c \in \mathfrak{R}$ und alle α, β und γ aus H genügt. Wir betrachten einen Vektorraum R^+ über $\mathfrak{R}$ mit einer Basis $\{\omega_\alpha\}_{\alpha \in H}$ und erhalten ersichtlich eine Strukturfunktion gemäß

$$F(a, b, \alpha, \beta, \gamma) = \begin{cases} a\ \alpha b\ c_{\alpha,\beta} & \text{für} \quad \alpha\beta = \gamma \\ 0 & \text{für} \quad \alpha\beta \neq \gamma, \end{cases}$$

die der Multiplikation der Monome

$$a\omega_\alpha b\omega_\beta = a\ \alpha b\ c_{\alpha,\beta}\ \omega_{\alpha\beta}$$

entspricht. Nach Satz 1 entsteht so ein Vektorraumring R über $\mathfrak{R}$, und wir wollen jeden Ring dieser Art ein *verschränktes Produkt des Ringes $\mathfrak{R}$ mit der Halbgruppe H* nennen. Diese Ringe sind assoziativ, da die Bedingung (8) von Satz 2 die Form

$$a\ \alpha b\ c_{\alpha,\beta}(\alpha\beta)c\ c_{\alpha\beta,\gamma} = a\ \alpha(b\ \beta c\ c_{\beta,\gamma})\ c_{\alpha,\beta\gamma}$$

annimmt, wofür die oben geforderten Relationen ersichtlich hinreichend (aber im allgemeinen nicht notwendig) sind. Auf diese Weise erhält man verschränkte Produkte von $\mathfrak{R}$ mit H, wie sie Bódi in [2] definiert und untersucht, wenn man statt Endomorphismen nur Automorphismen von $\mathfrak{R}$ zuläßt und noch fordert, daß $\mathfrak{R}$ ein Einselement hat und die Elemente $c_{\alpha,\beta} \in \mathfrak{R}$ Inverse besitzen. Wählt man insbesondere $\mathfrak{R}$ als Schiefkörper, H als Gruppe und die Funktionensysteme so, daß sie eine Schreiersche Gruppenerweiterung $\mathfrak{R}^* \circ H$ mit der multiplikativen Gruppe $\mathfrak{R}^*$ von $\mathfrak{R}$ als Normalteiler liefern, so entstehen die von Rédei in [5] eingeführten verschränkten Produkte.

Wir bemerken bereits hier, daß alle diese Beispiele von Vektorraumringen im allgemeinen keine Algebren i. w. S. sind; die speziellen Oreschen Polynomringe und die verschränkten Produkte stellen auch zugleich Beispiele für monomiale Ringe dar.

§ 2. Algebren im weiteren Sinne

Es sei R ein Vektorraumring über $\mathfrak{R}$ und $\{\omega_i\}$ eine Basis von R über $\mathfrak{R}$. Wir nehmen zunächst an, daß $\mathfrak{R}$ ein Einselement e besitzt; dann legen wegen[4]) $e\omega_i = \omega_i$

[4]) Das Einselement e von $\mathfrak{R}$ ist stets Einheitsoperator für alle Elemente aus R.

die Werte $F(e, e, i, j, k)$ der Strukturfunktion gerade die Multiplikation der Basiselemente fest. Von besonders einfacher Struktur werden dann diejenigen Vektorraumringe R sein, für die mit den Produkten der Elemente einer (geeigneten) Basis bereits die gesamte Multiplikation in R festgelegt ist, und zwar in einer Weise, die keinerlei weitere Kenntnisse über R und $\mathfrak{R}$ benötigt. Dafür liegt schon im Hinblick auf (6) der Ansatz

$$F(a, b, i, j, k) = a \cdot b \cdot F(e, e, i, j, k) \tag{9}$$

nahe. Gehen wir dann gemäß $F(e, e, i, j, k) = c_{ij}^{(k)}$ zur geläufigen Schreibweise mit Strukturkonstanten über, so erhalten wir an Stelle von (4) und (5)

$$a\omega_i b\omega_j = ab \sum_k c_{ij}^{(k)} \omega_k = (ab)(\omega_i \omega_j) \tag{10}$$

bzw.

$$\left(\sum a_i \omega_i\right)\left(\sum b_j \omega_j\right) = \sum_{i,j,k} a_i b_j c_{ij}^{(k)} \omega_k = \sum_{i,j} (a_i b_j)(\omega_i \omega_j). \tag{11}$$

Auf diese Weise gelangen wir (abgesehen von der hier nicht geforderten Assoziativität der Multiplikation) zu dem von Pickert in [4] eingeführten Begriff der *Algebra im weiteren Sinne* als einem Vektorraumring R über $\mathfrak{R}$, dessen Multiplikation mit einer geeigneten Basis $\{\omega_i\}$ durch (10) oder (11) festgelegt wird. Verwenden wir wie Pickert nur die linken Gleichungen von (10) bzw. (11), so wird die Existenz eines Einselementes von $\mathfrak{R}$ nicht benötigt, womit der Begriff der Algebra i. w. S. auch für Grundringe ohne Einselement erklärt ist. Wir können auch diesen Fall unseren Betrachtungen in § 1 unterordnen, wenn wir statt (9) mit einem beliebigen Rechtseinselement e_r von $\mathfrak{R}$

$$F(a, b, i, j, k) = a \cdot b \cdot F(e_r, e_r, i, j, k) = a \cdot b \cdot c_{ij}^{(k)} \tag{9'}$$

schreiben. Freilich legen dann die Strukturkonstanten $F(e_r, e_r, i, j, k) = c_{ij}^{(k)}$ nicht die Multiplikation der Basiselemente ω_i selbst, sondern die der Elemente $e_r\omega_i$ fest, und wir erhalten analog zu (10) und (11)

$$a\omega_i b\omega_j = ab \sum_k c_{ij}^{(k)} \omega_k = (ab)(e_r \omega_i e_r \omega_j) \tag{10'}$$

bzw.

$$\left(\sum a_i \omega_i\right)\left(\sum b_j \omega_j\right) = \sum_{i,j,k} a_i b_j c_{ij}^{(k)} \omega_k = \sum_{i,j} (a_i b_j)(e_r \omega_i e_r \omega_j). \tag{11'}$$

Jedoch empfiehlt es sich, auch dann mit den Relationen (10) und (11) weiterzuarbeiten, was stets möglich ist, wenn man von vornherein von der Basis $\{\omega_i\}$ zu der Basis $\{e_r\omega_i\}$ übergeht und diese wieder mit $\{\omega_i\}$ bezeichnet, womit man erreicht, daß das (willkürlich ausgewählte) Rechtseinselement e_r von $\mathfrak{R}$ auf die Basiselemente als Einheitsoperator wirkt. Wir wollen eine Basis $\{\omega_i\}$ einer Algebra i. w. S. über $\mathfrak{R}$, die diese Eigenschaft hat und für die (10) bzw. (11) gilt, eine *Algebrenbasis* von R über $\mathfrak{R}$ (i. w. S.) nennen.

Satz 3. *Ist R eine Algebra i. w. S. über $\mathfrak{R}$ mit der Algebrenbasis $\{\omega_i\}_{i \in I}$, so wird durch die Produkte der Basiselemente*

$$\omega_i \omega_j = \sum_k c_{ij}^{(k)} \omega_k$$

ein System von Strukturkonstanten $c_{ij}^{(k)}$ *gegeben, welches die Multiplikation in* R *gemäß* (11) *festlegt. Die Algebra* R *ist also durch* $\mathfrak{R}$, *die Mächtigkeit von* I, *die oben angegebene Auszeichnung eines Rechtseinselementes* e_r *von* $\mathfrak{R}$ *und die Strukturkonstanten* $c_{ij}^{(k)}$ *bis auf Isomorphie eindeutig bestimmt. Umgekehrt entsteht zu jedem Ring* $\mathfrak{R}$, *einer Indexmenge* I, *einem willkürlich ausgezeichneten Rechtseinselement* e_r *von* $\mathfrak{R}$ *und willkürlich vorgegebenen Strukturkonstanten*[5] $c_{ij}^{(k)} \in \mathfrak{R}$ *eine Algebra* R *i. w. S. mit diesen Strukturkonstanten, indem man in einem* $\mathfrak{R}$*-Vektorraum* R^+ *mit einer Basis* $\{\omega_i\}_{i\in I}$ *eine Multiplikation durch* (11) *einführt.*

Der Beweis ergibt sich sofort, wenn man gemäß (9′) zu einer Strukturfunktion übergeht und Satz 1 anwendet. Weiterhin erhalten wir aus Satz 2 unmittelbar:

Satz 4. *Eine Algebra* R *i. w. S. über* $\mathfrak{R}$ *ist genau dann assoziativ, wenn ihre Strukturkonstanten den Assoziativitätsbedingungen*

$$\sum_t c_{ij}^{(t)} c c_{tk}^{(l)} = c \sum_t c_{jk}^{(t)} c_{it}^{(l)} \tag{12}$$

für alle $c \in \mathfrak{R}$ *und alle* i, j, k *und* l *aus* I *erfüllen. Dagegen ist für die Assoziativität der Multiplikation der Basiselemente bereits*

$$\sum_t c_{ij}^{(t)} c_{tk}^{(l)} = \sum_t c_{jk}^{(t)} c_{it}^{(l)} \tag{13}$$

für alle i, j, k *und* l *notwendig und hinreichend.*

Zur Unterstützung der damit vorgenommenen Auszeichnung der Algebren i. w. S. unter allen Vektorraumringen bemerken wir noch, daß zu jedem Vektorraumring R über einem Ring $\mathfrak{R}$ mit Einselement[6]) eine Algebra $\tilde{R}$ i. w. S. mit der gleichen Multiplikation der Basiselemente korrespondiert. Man braucht ja nur von der Strukturfunktion $F(a, b, i, j, k)$ von R gemäß

$$\tilde{F}(a, b, i, j, k) = a \cdot b \cdot F(e, e, i, j, k)$$

zu einer neuen Strukturfunktion und damit zu $\tilde{R}$ überzugehen. Allerdings überträgt sich dabei die Assoziativität von R auf $\tilde{R}$ im allgemeinen nur dann, wenn die als Strukturkonstante von $\tilde{R}$ verwendeten Werte $F(e, e, i, j, k)$ der Strukturfunktion von R im Zentrum von $\mathfrak{R}$ liegen, wie aus Satz 4 hervorgeht.

Aus dieser Überlegung ergibt sich auch, daß ein Orescher Polynomring R nur in dem trivialen Fall eine Algebra i. w. S. (natürlich bezüglich der gleichen Basis $\{x^i\}$) ist, wenn für η die identische Abbildung und für δ der Nullendomorphismus gewählt werden, also der übliche Polynomring $\mathfrak{R}[x]$ vorliegt. Wie man sich leicht überlegt korrespondiert nämlich sonst in dem eben beschriebenen Sinne zu R gerade $\tilde{R} = \mathfrak{R}[x]$ als Algebra i. w. S. mit gleicher Multiplikation der Basiselemente, aber nicht gleicher Multiplikation in R bzw. $\tilde{R}$, womit aus Satz 3 folgt, daß nur $\tilde{R}$ Algebra i. w. S. sein kann. Entsprechend stellt man fest, daß ein verschränktes Produkt

[5]) Es versteht sich, daß natürlich bei festem i und j nur endlich viele $c_{ij}^{(k)} \neq 0$ gewählt werden dürfen.

[6]) Auch hier könnte man sich wie oben leicht von dieser Einschränkung frei machen.

genau dann Algebra i. w. S. ist, wenn die zugehörigen Endomorphismensysteme trivial gewählt werden, also $\alpha b = b$ für alle α und b gilt.

Für eine nähere Untersuchung der Algebren i. w. S. verweisen wir auf [4]. Wir bemerken nur noch, daß auch für nicht notwendig assoziative Algebren R i. w. S. über $\mathfrak{R}$ die Existenz eines Linkseinselementes von R bereits die Existenz eines Einselementes von $\mathfrak{R}$ und die Möglichkeit der isomorphen Einbettung von $\mathfrak{R}$ in R nach sich zieht. Schließlich heben wir hinsichtlich des Zusammenhanges der Algebren i. w. S. mit den üblicher Weise betrachteten Algebren noch folgendes hervor:

Sehen wir von der Verwendung der Strukturkonstanten einmal ab, so haben wir diejenigen Vektorraumringe R über $\mathfrak{R}$ als Algebren i. w. S. ausgezeichnet, die bezüglich einer geeigneten Basis $\{\omega_i\}$ stets

$$(10)\qquad (a\omega_i)(b\omega_j) = (ab)(\omega_i\omega_j)$$

oder, was auf das gleiche hinausläuft,

$$(11)\qquad (\sum a_i\omega_i)(\sum b_j\omega_j) = \sum_{i,j} (a_ib_j)(\omega_i\omega_j)$$

erfüllen. Wie man leicht nachprüft, folgt daraus (unter Verwendung von $e_r\omega_i = \omega_i$)

$$(14)\qquad a(\omega_i\omega_j) = (a\omega_i)\omega_j = \omega_i(a\omega_j)$$

und

$$(15)\qquad a(\alpha\beta) = (a\alpha)\beta \quad \text{für } \alpha\in R,\ \beta\in R,$$

während umgekehrt (14) und (15) wiederum (10) und damit (11) nach sich zieht[7]). Man kann aber leicht Beispiele dafür angeben, daß bei einem Wechsel der Basis die Eigenschaften (10), (11) und (14) nicht erhalten bleiben. So können wir nach den Sätzen 3 und 4 etwa eine (assoziative) Algebra R i. w. S. über dem Quaternionenkörper $\mathfrak{R}$ (mit den Quaternioneneinheiten i, j, k) durch die Strukturtafel der Basiselemente

$$\begin{array}{c|cc} & \omega_1 & \omega_2 \\ \hline \omega_1 & \omega_1 & \omega_2 \\ \omega_2 & \omega_2 & \omega_1 \end{array}$$

und (11) definieren. Gehen wir dann zu der Basis $\mu_1 = \omega_1$ und $\mu_2 = i\omega_2$ über, so gilt z. B. statt (10), (11) bzw. (14_2)

$$(j\mu_2)\mu_2 = (ji\omega_2)i\omega_2 = j\cdot i^2\omega_1 = -j\omega_1 \neq$$
$$\neq \mu_2(j\mu_2) = (i\omega_2)(ji\omega_2) = iji\omega_1 = j\omega_1\,.$$

Dies zeigt, daß der Begriff der Algebra i. w. S. in der Tat wesentlich von der jeweils zu Grunde gelegten Basis abhängig ist. Dagegen gelten in den üblicher Weise durch

$$(2)\qquad a(\alpha\beta) = (a\alpha)\beta = \alpha(a\beta)$$

definierten Algebren, die wir mit [4] auch Algebren im engeren Sinne (i. e. S.) nennen

[7]) Die folgende Überlegung zeigt auch, daß man nicht etwa allein von (15) auf (10) bzw. (11) schließen kann.

wollen, die Regeln (10), (11) und (14) für jede Basis $\{\omega_i\}$ von R über $\mathfrak{R}$. Allerdings ist eben diese Forderung (2) sehr einschneidend. Bekanntlich ergibt sich aus ihr

$$(ab-ba)(\alpha\beta) = o$$

für alle a, b, α und β. Wäre dann $\mathfrak{R}$ nicht kommutativ, so dürfte das von allen $\alpha\beta$ erzeugte Ideal R^2 von R kein über $\mathfrak{R}$ linear unabhängiges Element enthalten; insbesondere müßte für jede Vektorraumbasis $\{\omega_i\}$ von R^+ über $\mathfrak{R}$ stets $\{\omega_i\} \cap R^2 = \emptyset$ gelten. Schon die Existenz eines Rechts- oder Linkseinselementes in R (vgl. [4]) führt dann wegen $R^2 = R$ zur Kommutativität von $\mathfrak{R}$. Umgekehrt fallen natürlich über einem kommutativen Grundring $\mathfrak{R}$ die Begriffe der Algebra i. w. S. und i. e. S. zusammen.

§ 3. Doppelalgebren

Als nächstes wollen wir den von RÉDEI geprägten Begriff der Doppelalgebra (vgl. [5], § 64) in unsere allgemeinen Zusammenhänge einordnen. Während RÉDEI die Theorie der Algebren auf Algebren i. e. S. und damit im wesentlichen auf kommutative Grundringe beschränkt, führt er zur Erfassung entsprechender Strukturen über nichtkommutativen Grundringen den Begriff der Doppelalgebra ein. Dazu sei R ein $\mathfrak{R}$-Doppelmodul, d. h. sowohl ein $\mathfrak{R}$-Links- wie $\mathfrak{R}$-Rechtsmodul, der noch

(16) $$(a\alpha)b = a(\alpha b)$$

für alle $a \in \mathfrak{R}$, $b \in \mathfrak{R}$ und $\alpha \in R$ erfüllt. Insbesondere heißt R ein $\mathfrak{R}$-Doppelvektorraum, wenn eine Basis $\{\omega_i\}$ von R über $\mathfrak{R}$ als Linksvektorraum existiert, für deren Elemente überdies

(17) $$a\omega_i = \omega_i a$$

für alle $a \in \mathfrak{R}$ gilt[8]). Schließlich heißt ein assoziativer Ring R eine $\mathfrak{R}$-Doppelalgebra, wenn R ein Doppelvektorraum über R bezüglich der Basis $\{\omega_i\}$ ist und stets die Relationen

(18) $$(\alpha a)\beta = \alpha(a\beta)$$

sowie

(14) $$a(\omega_i\omega_j) = (a\omega_i)\omega_j = \omega_i(a\omega_j)$$

gelten. Natürlich ist auch diese Begriffsbildung von der Basis $\{\omega_i\}$ abhängig. Wir zeigen sogleich allgemein:

Satz 5. *Jede $\mathfrak{R}$-Doppelalgebra R ist bezüglich der gleichen Basis $\{\omega_i\}$ auch (linksseitige) Algebra i. w. S. über $\mathfrak{R}$. Umgekehrt kann jede Algebra R i. w. S. über einem Ring $\mathfrak{R}$ mit Einselement e bezüglich der gleichen Basis $\{\omega_i\}$ zur $\mathfrak{R}$-Doppelalgebra gemacht werden, wenn man die Rechtsoperatoranwendung von $a \in \mathfrak{R}$ auf $\alpha = \Sigma a_i\omega_i \in R$ wie folgt definiert:*

(19) $$\alpha a = (\Sigma a_i\omega_i)a = \Sigma a_i a\omega_i.$$

[8]) Doppelvektorräume können also nur über Grundringen mit Einselement existieren.

Beweis. Unter Verwendung von (18), (16), (17), (1) und (14) erhalten wir

$$(a\omega_i)(b\omega_j)=((a\omega_i)b)\omega_j=(a(\omega_i b))\omega_j=(a(b\omega_i))\omega_j$$
$$=((ab)\omega_i)\omega_j=(ab)(\omega_i\omega_j),$$

womit jede $\mathfrak{R}$-Doppelalgebra auch (10) erfüllt und damit Algebra i. w. S. ist. Für die Umkehrung stellen wir zunächst fest, daß durch (19) eine eindeutig bestimmte Operatoranwendung erklärt wird (die natürlich auch von der zugrunde gelegten Basis abhängig ist). Wie man leicht nachrechnet, wird damit R auch zu einem $\mathfrak{R}$-Rechtsmodul, und es gilt (16) und (18), letzteres z. B. gemäß:

$$(\alpha a)\beta=(\Sigma(a_i a)\omega_i)(\Sigma b_j\omega_j)=\Sigma(a_i a)b_j(\omega_i\omega_j)$$
$$=\Sigma a_i(ab_j)(\omega_i\omega_j)=(\Sigma a_i\omega_i)(\Sigma ab_j\omega_j)=\alpha(a\beta).$$

Schließlich benötigt man zum Nachweis von (17) das Einselement e von $\mathfrak{R}$:

$$\omega_i a=(e\omega_i)a=(ea)\omega_i=a\omega_i.$$

Wir bemerken noch, daß sowohl die Begriffsbildung der $\mathfrak{R}$-Doppelalgebra R wie auch unser Satz nicht von der (in [5] vorausgesetzten) Assoziativität von R abhängen. Weiterhin lehrt unser Beweis, daß zur Festlegung des Begriffes der Doppelalgebra die Forderung (16) entbehrlich ist, da von ihr nur $(a\omega_i)b=a(\omega_i b)$ verwendet wurde, was aber bereits aus (17), (1) und den entsprechenden Linksmodulgesetzen folgt, und daß man sich bei (14) auf die erste Gleichung beschränken könnte. Doch läuft unseres Erachtens Satz 5 gerade darauf hinaus, die Heranziehung der Rechtsoperatoranwendung überhaupt als unnötig anzusehen und den Begriff der Doppelalgebra durch den strukturell einfacheren der Algebra i. w. S. zu ersetzen.

§ 4. Monomiale Ringe

Der von Rédei in [5], § 66 intendierte allgemeine Begriff des *monomialen Ringes R über einem Ring $\mathfrak{R}$* läßt sich mit Hilfe der Überlegungen aus § 1 wie folgt fassen: R ist ein Vektorraumring über $\mathfrak{R}$, der eine solche Basis $\{\omega_i\}$ besitzt, daß die zugehörige Strukturfunktion $F(a, b, i, j, k)$ bei festen Argumenten a, b, i und j höchstens für ein k einen von 0 verschiedenen Wert annimmt. Wir werden dann auch diese Basis bzw. die zugehörige Strukturfunktion monomial nennen; es liegt auf der Hand, daß bei einem Basiswechsel eine monomiale Basis nicht wieder in eine solche übergehen muß. Beispiele von Vektorraumringen, die in dem eben definierten Sinne (assoziative) monomiale Ringe sind, haben wir bereits in § 1 gegeben.

In Anlehnung an Rédei bezeichnen wir die Elemente einer monomialen Basis von R über $\mathfrak{R}$ mit $\omega_A, \omega_B, \ldots$, wobei die Indices $A, B, \ldots$ eine Indexmenge H durchlaufen. Um über die Sätze von § 1 hinausgehende Aussagen machen zu können, setzen wir noch voraus, daß das Verschwinden der monomialen Strukturfunktion

$$F(a, b, A, B, C) \tag{20}$$

für $a\neq 0$ und $b\neq 0$ nur von den Argumenten A, B und C abhängt. Dies trifft etwa für die eben zitierten Beispiele weitgehend zu. Auch ist diese Voraussetzung stets erfüllt, wenn die monomiale Basis zugleich Algebrenbasis i. w. S. ist[9]), wie aus

$$F(a, b, A, B, C) = abF(e_r, e_r, A, B, C)$$

folgt; doch wollen wir auf diesen Fall erst später zu sprechen kommen.

Im folgenden verstehen wir also unter einem monomialen Ring R über $\mathfrak{R}$ einen solchen, der eine monomiale Basis $\{\omega_A, \omega_B, \ldots\}$ besitzt, wo die zugehörige Strukturfunktion auch der bei (20) formulierten Bedingung genügt. Nur in diesem Sinne wollen wir von jetzt an auch die Bezeichnungen „monomiale Basis" und „monomiale Strukturfunktion" verwenden. Das bedeutet also, daß in einem monomialen Ring R das Produkt zweier Monome $a\omega_A \neq o$ und $b\omega_B \neq o$ bei festen ω_A und ω_B entweder stets verschwindet oder stets ein Monom $F(a, b, A, B, C)\omega_C \neq o$ mit festem ω_C ergibt. Falls es nun zu jedem Indexpaar A, B einen Index C mit $F(a, b, A, B, C) \neq 0$ gibt, so wird durch $AB = C$ in der Indexmenge H eine Multiplikation erklärt. Sonst erweitern wir die Indexmenge H durch ein nicht in H vorhandenes Element $\mathcal{O}$ zur Menge $H_0 = H \cup \{\mathcal{O}\}$ und erklären in H_0 eine Multiplikation wie folgt:

(21) $AB = C$ falls $F(a, b, A, B, C) \neq 0$ für $a \neq 0$ und $b \neq 0$,
$AB = \mathcal{O}$ falls $F(a, b, A, B, X) = 0$ für $a \neq 0$, $b \neq 0$ und alle $X \in H$,
$A\mathcal{O} = \mathcal{O}A = \mathcal{O}$.

Entsprechend erweitern wir auch den Definitionsbereich der Strukturfunktion $F(a, b, A, B, C)$, indem wir $F(a, b, A, B, C) = 0$ setzen, falls für A oder B oder C das Element $\mathcal{O}$ eingesetzt wird. Deuten wir noch $\omega_{\mathcal{O}}$ in $0\omega_{\mathcal{O}}$ als ein beliebiges Basiselement, so gilt für die Multiplikation in R ganz allgemein

(22) $$a\omega_A b\omega_B = F(a, b, A, B, AB)\omega_{AB}.$$

Mit diesen Vorbereitungen gelangen wir zu folgendem Satz, wobei wir daran erinnern, daß nach Satz 1 aus jedem Vektorraum R^+ über $\mathfrak{R}$ mit einer Basis $\{\omega_A, \omega_B, \ldots\}$ durch eine monomiale Strukturfunktion $F(a, b, A, B, C)$ ein monomialer Ring R mit dieser Basis entsteht.

Satz 6. *Es sei R ein monomialer Ring über $\mathfrak{R}$ mit einer monomialen Basis $\{\omega_A, \omega_B, \ldots\}$ und der monomialen Strukturfunktion $F(a, b, A, B, C)$. Dann ist R genau dann assoziativ, wenn die Indexmenge H bzw. die erweiterte Indexmenge H_0 mit der eben erklärten Multiplikation eine Halbgruppe bildet und die (gegebenenfalls für Argumente aus H_0 erweiterte) Strukturfunktion die Bedingungen*

(23) $$F(F(a, b, A, B, AB), c, AB, C, (AB)C) = \\ = F(a, F(b, c, B, C, BC), A, BC, A(BC))$$

für alle a, b und c aus $\mathfrak{R}$ und alle A, B und C aus H erfüllt.

Beweis. Wir zeigen die Behauptung sogleich für den Fall, daß in R Monome $a\omega_A \neq o$ und $b\omega_B \neq o$ mit $a\omega_A b\omega_B = o$ auftreten und mit der erweiterten Indexmenge H_0 gearbeitet werden muß; der andere Fall ergibt sich (unter entsprechenden Vereinfachungen) auf die gleiche Weise.

[9]) Nur dieser Fall wird bei Rédei a. a. O. implizit betrachtet.

Ist R assoziativ, dann bildet die Menge $\mathfrak{H}$ aller Monome $x\omega_A$ mit beliebigen $x \neq 0$ aus $\mathfrak{R}$ und $A \in H$ sowie das Element $o = 0\omega_X$ für alle $X \in H$ mit der Multiplikation von R eine Halbgruppe. Durch

$$x\omega_A \to A \qquad \text{für alle } x \neq 0 \text{ aus } \mathfrak{R},$$

$$o = 0\omega_X \to \mathcal{O}$$

wird $\mathfrak{H}$ ersichtlich eindeutig und auch relationstreu auf H_0 abgebildet. Letzteres folgt aus

$$x\omega_A y\omega_B = F(x, y, A, B, C)\omega_C \text{ für } AB = C,$$

$$x\omega_A y\omega_B = o \qquad \text{für } AB = \mathcal{O},$$

sowie

$$0\omega_A y\omega_B = x\omega_A 0\omega_B = 0\omega_A 0\omega_B = 0\omega_X.$$

Als homomorphes Bild der Halbgruppe $\mathfrak{H}$ ist H_0 eine Halbgruppe. Weiterhin ist für die Assoziativität von R gemäß Satz 2 notwendig und hinreichend, daß die Strukturfunktion (8), also

$$(24) \qquad \sum_T F\big(F(a, b, A, B, T), c, T, C, L\big) = \sum_T F\big(a, F(b, c, B, C, T), A, T, L\big)$$

für alle a, b und c aus $\mathfrak{R}$ und alle A, B, C und L aus H erfüllt, wobei aber eben hier für diese Argumente wie auch für T nur Elemente aus H zugelassen sind. Nun ist in (24) die linke Summe gleich

$$F\big(F(a, b, A, B, AB), c, AB, C, (AB)C\big) \quad \text{falls } AB \in H \text{ und } (AB)C \in H,$$

während sonst sogar alle Summanden verschwinden. Entsprechendes gilt für die rechte Summe von (24) mit

$$F\big(a, F(b, c, B, C, BC), A, BC, A(BC)\big) \quad \text{falls } BC \in H \text{ und } A(BC) \in H.$$

Führen nun alle Produkte AB, $(AB)C$, BC, $A(BC)$ nicht aus H heraus, so reduziert sich (24) gerade auf (23). Andernfalls gilt aber stets $(AB)C = A(BC) = \mathcal{O}$, da wir nach dem bereits bewiesenen nur noch den Fall zu berücksichtigen brauchen, daß H_0 Halbgruppe ist. Dann verschwinden aber bei (24) wie bei (23) beide Seiten, das erstere nach der eben getroffenen Feststellung, das letztere gemäß unserer Vereinbarung über das (bei (23) zugelassene) Auftreten des Argumentes $\mathcal{O}$ in der Strukturfunktion.

Bisher haben wir in der Indexmenge H bzw. in H_0 eine Multiplikation eingeführt, die diesen Mengen durch R bzw. eine vorgegebene Strukturfunktion aufgeprägt wurden. Wir können aber auch von einer in $H = \{A, B, \ldots\}$ vorgegebenen Multiplikation ausgehen und solche monomialen Ringe R über $\mathfrak{R}$ mit einer Basis $\{\omega_A, \omega_B, \ldots\}$ betrachten, für welche die Multiplikation von R mit der Multiplikation von H im Einklang steht, d. h.

$$(25) \qquad \text{für } AB \neq C \text{ stets } F(a, b, A, B, C) = 0,$$

$$\text{für } AB = C \text{ entweder stets } F(a, b, A, B, C) = 0$$

$$\text{oder } F(a, b, A, B, C) \neq 0 \text{ für alle } a \neq 0 \text{ und } b \neq 0$$

gilt. Ist insbesondere H eine Halbgruppe und der monomiale Ring R assoziativ, so nennen wir ihn einen *verallgemeinerten Halbgruppenring* von H über $\mathfrak{R}$ [10]).

Satz 7. *Ein monomialer Ring R über $\mathfrak{R}$ mit einer monomialen Basis $\{\omega_A, \omega_B, \ldots\}$, deren Indices eine Halbgruppe H durchlaufen, ist genau dann ein verallgemeinerter Halbgruppenring von H über $\mathfrak{R}$, wenn die zugehörige Strukturfunktion $F(a, b, A, B, C)$ gemäß* (25) *mit der Multiplikation von H im Einklang steht und der Assoziativitätsbedingung* (23) *von Satz* 6 *genügt.*

Der allein erforderliche Nachweis für die Assoziativität von R ergibt sich mit dem gleichen Gedankengang wie im Beweis zu Satz 6.

Wir betrachten nun einen assoziativen, monomialen Ring R über $\mathfrak{R}$, der den Voraussetzungen von Satz 6 genügt. Für den Fall, daß das Produkt zweier von o verschiedener Monome $a\omega_A$ und $b\omega_B$ von R stets $a\omega_A b\omega_B \neq o$ erfüllt, wird die Indexmenge H durch die bei (21) erklärte Multiplikation zur Halbgruppe, und R ist ersichtlich ein verallgemeinerter Halbgruppenring von H über $\mathfrak{R}$ [11]). Falls dagegen Produkte von o verschiedener Monome aus R verschwinden, bildet nur die erweiterte Indexmenge H_0 mit der Multiplikation (21) eine Halbgruppe, in der $\mathcal{O}$ Nullelement ist. Dann können wir aber einen verallgemeinerten Halbgruppenring R' von H_0 über $\mathfrak{R}$ bilden, indem wir $\{\omega_{\mathcal{O}}, \omega_A, \omega_B, \ldots\}$ als Basis von R' über $\mathfrak{R}$ nehmen und eine Multiplikation durch die vorn bereits auf den Definitionsbereich H_0 erweiterte Strukturfunktion $F(a, b, A, B, C)$ von R einführen [12]). In der Tat ist $F(a, b, A, B, C)$ dann auch bezüglich der Basis $\{\omega_{\mathcal{O}}, \omega_A, \omega_B, \ldots\}$ eine monomiale Strukturfunktion, während die Assoziativitätsbedingung (23) gerade in Satz 6 nachgewiesen wurde. Ersichtlich bilden dann die Monome $x\omega_{\mathcal{O}}$ für alle $x \in \mathfrak{R}$ als Annullatoren ein zulässiges Ideal $(e_r \omega_{\mathcal{O}})$ von R', und es gilt $R \cong R'/(e_r \omega_{\mathcal{O}})$. Dabei besteht die Restklassenbildung anschaulich gesprochen einfach darin, das Basiselement $\omega_{\mathcal{O}}$ mit dem Nullelement von R' zu identifizieren. Wir fassen zusammen:

Satz 8. *Ein assoziativer monomialer Ring R über $\mathfrak{R}$ mit der Basis $\{\omega_A\}_{A \in H}$ ist bezüglich der bei* (21) *erklärten Multiplikation in der Indexmenge H bzw. H_0 entweder ein verallgemeinerter Halbgruppenring von H über $\mathfrak{R}$ oder isomorph zum Restklassenring eines verallgemeinerten Halbgruppenringes R' von H_0 über $\mathfrak{R}$ nach dem Ideal $(e_r \omega_{\mathcal{O}})$ von R'.*

Wir wenden unsere Ergebnisse nun auf den Fall an, daß eine monomiale Basis $\{\omega_A, \omega_B, \ldots\}$ von R über $\mathfrak{R}$ existiert, die zugleich Algebrenbasis i. w. S. ist; wir nennen dann R eine monomiale Algebra i. w. S. über $\mathfrak{R}$. Die zugehörigen Strukturfunktionen sind durch

$$F(a, b, A, B, C) = a \cdot b \cdot F(e_l, e_r, A, B, C) = a \cdot b \cdot c_{A,B}^{(C)} \tag{26}$$

---

[10]) Mit $F(a, b, A, B, C) = 0$ für $AB \neq C$ und $F(a, b, A, B, C) = ab$ für $AB = C$ entsteht so der Halbgruppenring von H über $\mathfrak{R}$ im üblichen Sinne.

[11]) Dieser Halbgruppenring hat die spezielle Eigenschaft, daß in (25) der erste Fall bei $AB = C$ nie eintritt.

[12]) Der Unterschied zu vorn besteht dann darin, daß wir $\omega_{\mathcal{O}}$ jetzt als ein weiteres Basiselement ansehen und nicht als eines der Basiselemente $\omega_A, \omega_B, \ldots$ deuten. Übrigens liegt auch hier ein Spezialfall von (25) vor, da wir für $AB = \mathcal{O}$ stets $F(a, b, A, B, \mathcal{O}) = 0$ haben.

und die Forderung gekennzeichnet, daß für jedes Paar A, B aus H höchstens ein $c_{A,B}^{(C)} \neq 0$ ist. Die Einführung einer Multiplikation in H bzw. H_0 erfolgt dann analog zu (21) gemäß

$$
\begin{aligned}
&AB = C \quad \text{falls} \quad c_{A,B}^{(C)} \neq 0, \\
(27) \qquad &AB = \mathcal{O} \quad \text{falls} \quad c_{A,B}^{(X)} = 0 \quad \text{für alle} \quad X \in H, \\
&A\mathcal{O} = \mathcal{O}A = \mathcal{O},
\end{aligned}
$$

und wir setzen auch hier $c_{A,B}^{(C)} = 0$, wenn für A oder B oder C das Element $\mathcal{O} \in H_0$ eingesetzt wird. Denkt man sich die Multiplikation von H bzw. H_0 bereits anderweitig fixiert, so können wir die Schreibweise durch

$$
(28) \qquad c_{A,B} = c_{A,B}^{(AB)}
$$

vereinfachen, da alle anderen Strukturkonstanten $c_{A,B}^{(C)}$ ohnehin verschwinden, während für die $c_{A,B}$

$$
(29) \qquad c_{A,B} = 0 \text{ genau dann, wenn } AB = \mathcal{O}
$$

gilt. An Stelle von (10) und (22) erhalten wir dann

$$
(30) \qquad a\omega_A b\omega_B = ab\omega_A\omega_B = ab\, c_{A,B}\omega_{AB},
$$

wobei natürlich wieder $\omega_{\mathcal{O}}$ in $0\omega_{\mathcal{O}}$ als ein beliebiges Basiselement zu deuten ist. Mit [5] nennen wir die $c_{A,B}$ in (28) das *Faktorensystem* von R über $\mathfrak{R}$ bezüglich $\{\omega_A\}_{A \in H}$, wobei wir nochmals betonen, daß dieser Begriffsbildung eine Festlegung der Multiplikation in H bzw. in H_0 vorauszugehen hat. Dann geht Satz 6 über in

Satz 9. *Es sei R eine monomiale Algebra i. w. S. über dem Ring $\mathfrak{R}$ mit einer monomialen Algebrenbasis $\{\omega_A, \omega_B, \ldots\}$ und dem Faktorensystem $\{c_{A,B}\}$. Dann ist R genau dann assoziativ, wenn die Indexmenge H bzw. die erweiterte Indexmenge H_0 mit der dem Faktorensystem $\{c_{A,B}\}$ zu Grunde gelegten Multiplikation eine Halbgruppe bildet und die Bedingung*

$$
(31) \qquad c_{A,B}\, c\, c_{AB,C} = c\, c_{B,C}\, c_{A,BC}
$$

für alle $c \in \mathfrak{R}$ und alls A, B und C aus H gilt. Falls dabei alle Faktoren $c_{A,B}$ im Zentrum von $\mathfrak{R}$ liegen, kann in (31) *das Element c auf beiden Seiten gestrichen werden*[13]).

Damit erhält man also unter Verwendung von Satz 3 sämtliche assoziativen monomialen Algebren R i. w. S. über $\mathfrak{R}$ mit einer Basis $\{\omega_A\}_{A \in H}$, indem man für H eine beliebige Halgbruppe oder die Menge der von $\mathcal{O}$ verschiedenen Elemente einer Halbgruppe H_0 mit Nullelement $\mathcal{O}$ nimmt und dazu alle Faktorensysteme $\{c_{A,B}\}$ bestimmt, welche die Bedingungen (29) und (31) erfüllen. Andererseits kann man natürlich auch zu einer beliebigen Halbgruppe H ein Faktorensystem $\{c_{A,B}\}$ betrachten, welches nur der Bedingung (31) genügt. Die Strukturfunktion

$$
F(a, b, A, B, C) = \begin{cases} a \cdot b \cdot c_{A,B} & \text{für} \quad AB = C \\ 0 & \text{für} \quad AB \neq C \end{cases}
$$

[13]) Vgl. auch [5], Satz 154. Die dort zu treffende Zusatzbedingung entfällt hier auf Grund unserer Erweiterung von H zu H_0.

erfüllt dann nämlich (25) und wegen (31) eben (23), so daß auf diese Weise nach Satz 7 ein verallgemeinerter Halbgruppenring R von H über $\mathfrak{R}$ mit der Basis $\{\omega_A\}_{A\in H}$ entsteht, die zugleich Algebrenbasis i. w. S. ist. Etwas allgemeiner als RÉDEI nennen wir dann R einen *Halbgruppenring mit Faktorensystem* über $\mathfrak{R}$ und fassen zusammen (vgl. [5], Satz 155):

Satz 10. *Ist H eine Halbgruppe, so ist der Halbgruppenring R über $\mathfrak{R}$ mit dem Faktorensystem $\{c_{A,B}\}$ genau dann definiert, wenn die Bedingung* (31) *für alle $c\in\mathfrak{R}$ und alle A, B und C aus H erfüllt ist.*

Analog übertragen wir Satz 8 und erhalten:

Satz 11. *Es sei R eine assoziative monomiale Algebra i. w. S. über $\mathfrak{R}$ mit der Basis $\{\omega_A\}_{A\in H}$ und dem Faktorensystem $\{c_{A,B}\}$. Dann ist R bezüglich der diesem Faktorensystem zu Grunde liegenden Multiplikation in H bzw. H_0 entweder der Halbgruppenring von H über $\mathfrak{R}$ mit diesem Faktorensystem $\{c_{A,B}\}$ oder isomorph zum Restklassenring des Halbgruppenringes R' von H_0 über $\mathfrak{R}$ mit diesem Faktorensystem* [14] *$\{c_{A,B}\}$ nach dem Ideal $(e,\omega_{\mathcal{O}})$ von R'.*

Da dieses Faktorensystem nach den Vorbetrachtungen zu Satz 9 der Bedingung (29) genügt, ist es mitunter vorteilhaft, beliebige Halbgruppenringe R über $\mathfrak{R}$ mit einem Faktorensystem $\{d_{A,B}\}$ (wobei die $d_{A,B}$ ja unabhängig von (29) gewählt werden können) als monomiale Algebren mit Hilfe von Satz 11 auf andere Halbgruppenringe zurückzuführen, für deren Faktorensystem dann (29) erfüllt sein muß.

Wir erläutern das letztere durch ein recht einfaches Beispiel: Für die Halbgruppe H wählen wir die zyklische Gruppe $H=\{X^0, X^1, X^2\}$ der Ordnung 3 und bestimmen etwa im Ring $\mathfrak{R}$ der ganzen Zahlen ein Faktorensystem gemäß

$$d_{X^i,X^j}=\begin{cases}1 & \text{für} \quad i+j\leqq 2\\ 0 & \text{für} \quad i+j>2.\end{cases}$$

Wie man sieht, ist (31) (nicht aber (29)) erfüllt, und es existiert nach Satz 10 der Halbgruppenring R von H über $\mathfrak{R}$ mit diesem Faktorensystem. Wir können aber H auch als Untermenge der von $\mathcal{O}$ verschiedenen Elemente einer Halbgruppe H_0 mit der Strukturtafel

	X^0	X^1	X^2	$\mathcal{O}$
X^0	X^0	X^1	X^2	$\mathcal{O}$
X^1	X^1	X^2	$\mathcal{O}$	$\mathcal{O}$
X^2	X^2	$\mathcal{O}$	$\mathcal{O}$	$\mathcal{O}$
$\mathcal{O}$	$\mathcal{O}$	$\mathcal{O}$	$\mathcal{O}$	$\mathcal{O}$

auffassen und ein Faktorensystem gemäß

$$c_{A,B}=\left.\begin{cases}1 & \text{falls} \quad AB\neq\mathcal{O}\\ 0 & \text{falls} \quad AB=\mathcal{O}\end{cases}\right\}\quad \text{für alle } A, B \text{ aus } H$$

[14]) Man beachte, daß wir bereits $c_{A,\mathcal{O}}=c_{\mathcal{O},A}=c_{\mathcal{O},\mathcal{O}}=0$ definiert haben.

im Einklang mit (29) einführen. Dann ist R isomorph zum Restklassenring des Halbgruppenringes R' von H_0 über $\mathfrak{R}$ mit $\{c_{A,B}\}$ als Faktorensystem nach dem Ideal $(\omega_{\mathfrak{S}})$.

Schließlich ist für das Aufsuchen aller möglichen Faktorensysteme zu gewissen Halbgruppen H bzw. H_0 noch folgendes Korollar nützlich:

Korollar. *Ist $\{c_{A,B}\}$ ein Faktorensystem einer assoziativen monomialen Algebra oder eines Halbgruppenringes R über $\mathfrak{R}$, so gilt das gleiche für $\{c'_{A,B}\}$ mit $c'_{A,B} = -c_{A,B}$.*

Literaturverzeichnis

[1] K. Asano, Über die Quotientenbildung von Schiefringen, *J. Math. Soc. Japan*, 1 (1949), 73—78.
[2] А. А. Бовди, О скрещенных произведениях полугруппы и кольца, *Доклады Акад. Наук. СССР*, 137 (1961), 1267—1269.
[3] O. Ore, Theory of non-commutative polynomials, *Ann. of Math.*, 34 (1933), 480—508.
[4] G. Pickert, Bemerkungen zum Algebrenbegriff, *Math. Ann.*, 120 (1947—49), 158—164.
[5] L. Rédei, *Algebra*. I (Leipzig, 1959).
[6] D. Tamari, On a certain classification of rings and semigroups, *Bull. Amer. Math. Soc.*, 54 (1948), 153—158.

(Eingegangen am 15. Mai 1964)

Numerischer Wertebereich und normale Dilatationen

Von STEFAN HILDEBRANDT in Mainz (Deutschland)

1. Wir wollen für beschränkte lineare Operatoren T in einem komplexen Hilbertraum $\mathfrak{H}$ den numerischen Wertebereich $W(T)$ betrachten, der durch

$$W(T) = \{\lambda: \lambda = (Tx, x), \|x\| = 1\}$$

definiert ist. Mit HALMOS bezeichnen wir einen Operator N, der in einem Hilbertraum $\mathfrak{K}$ definiert ist, welcher $\mathfrak{H}$ als Unterraum enthält, als eine Dilatation von T, falls

$$T = PNP$$

ist. Dabei bedeutet P den orthogonalen Projektor von $\mathfrak{K}$ auf $\mathfrak{H}$. Ferner heißt N starke Dilatation von T, falls sogar

$$T^n = PN^nP \quad \text{für alle } n = 1, 2, \ldots$$

gilt.

In [1] hat HALMOS die folgenden Ergebnisse bewiesen:

(1) Der Abschluß des numerischen Wertebereiches eines Operators in $\mathfrak{H}$ ist der Durchschnitt der Abschlüsse der numerischen Wertebereiche seiner normalen Dilatationen in $\mathfrak{K} = \mathfrak{H} \oplus \mathfrak{H} = 2\mathfrak{H}$.

(2) Der Abschluß des numerischen Wertebereiches jeder normalen Kontraktion T (d. h. $\|T\| \leqq 1$) ist der Durchschnitt der Abschlüsse der numerischen Wertebereiche seiner unitären Dilatationen in $\mathfrak{K} = 2\mathfrak{H}$.

Offengeblieben war unter anderem die Frage, ob (2) für jede Kontraktion richtig bleibt. Immerhin konnte HALMOS durch Kombination von (1) und (2) zeigen, daß (2) für alle Operatoren T mit $\|T\| \leqq \frac{1}{3}$ gilt, falls man alle unitären Dilatationen aus $\mathfrak{K} = 4\mathfrak{H}$ zuläßt.

Im folgenden wollen wir einige Sätze beweisen, die (1) und (2) verschärfen und ein erster Schritt in Richtung des obigen Problems sind, nämlich:

(I) Der Abschluß des numerischen Wertebereichs eines Operators auf $\mathfrak{H}$ ist der Durchschnitt der Abschlüsse der numerischen Wertebereiche seiner starken normalen Dilatationen N von der Form $N = \alpha U + \beta I$, wo U einen unitären Operator, I die Identität und α, β komplexe Zahlen bedeuten.

(II) Der Abschluß $\overline{W(T)}$ des numerischen Wertebereiches eines Operators T ist der Durchschnitt der Abschlüsse der numerischen Wertebereiche seiner unitären Dilatationen, falls sich $\overline{W(T)}$ als Durchschnitt von konvexen, im Einheitskreis

$\{z: |z| \leqq 1\}$ gelegenen Spektralmengen von T (im Sinne von v. NEUMANN) darstellen läßt.

Also gilt (II) insbesondere für Operatoren T, bei denen $\overline{W(T)}$ eine Spektralmenge ist, z. B. für die normalen und subnormalen Operatoren, wo sogar das Spektrum eine Spektralmenge ist. Falls $\mathfrak{H}$ endlichdimensional ist, ist unser Resultat insofern schwächer, als sich über die Dimension des Erweiterungsraumes $\mathfrak{K}$, in dem die Dilatationen betrachtet werden, keine Einschränkungen machen lassen.

Anders als HALMOS stützen wir uns zum Beweis von (I) und (II) auf die v. Neumannsche Theorie der Spektralmengen (vgl. [3] und [4]) und auf die Theorie der starken unitären und normalen Dilatationen (vgl. [5] und [6]).

Ich danke Herrn Prof. FOIAŞ für die Vereinfachung des Beweises für den folgenden Satz 1, die er mir freundlicherweise mitgeteilt hat, und Herrn Prof. SZ.-NAGY für den Hinweis auf den Gegenstand.

2. Eine abgeschlossene Menge X der komplexen Ebene heißt Spektralmenge des beschränkten linearen Operators T, wenn $\sigma(T) \subseteqq X$ und wenn für jede rationale Funktion $f(\lambda)$ ohne Pole in X (für die dann nach dem Dunfordschen Funktionalkalkül der Operator $f(T)$ definiert ist) die Ungleichung

$$\|f(T)\| \leqq \|f\|_X \overset{\text{def}}{=} \sup \{|f(\lambda)|: \lambda \in X\}$$

erfüllt ist.

Satz 1. *Der Abschluß $\overline{W(T)}$ des numerischen Wertebereiches eines beliebigen beschränkten linearen Operators T in $\mathfrak{H}$ ist der Durchschnitt von Kreisscheiben, die sämtlich Spektralmengen von T sind.*

Beweis. Bekanntlich (s. [4], S. 423) ist die Halbebene $\operatorname{Re} \lambda \geqq 0$ eine Spektralmenge eines Operators T genau dann, wenn $\operatorname{Re}(Tx, x) \geqq 0$ für jeden Vektor x gilt. Daraus sieht man leicht, daß jede konvexe Spektralmenge den numerischen Wertebereich enthält. Daher brauchen wir nur zu zeigen, daß es zu jedem nicht in $\overline{W(T)}$ gelegenen Punkt z_0 eine Kreisscheibe $C_{\gamma,r} = \{z: |z-\gamma| \leqq r\}$ gibt, die ebenfalls z_0 nicht enthält und die Spektralmenge von T ist. Da $\overline{W(T)}$ nach einem Satz von TOEPLITZ und HAUSDORFF konvex ist, gibt es eine Stützgerade g an $\overline{W(T)}$ mit Stützpunkt $s \in \partial W(T)$[1]) derart, daß g den Punkt z_0 von $\overline{W(T)}$ trennt und die Verbindungsgerade h der Punkte z_0 und s auf g senkrecht steht. Da für lineare Abbildungen $p(z) = \alpha z + \beta$ die Beziehung $W(p(T)) = p(W(T))$ gilt und da durch $p(z)$ Spektralmengen in Spektralmengen übergeführt werden, kann man annehmen, daß h mit der reellen und g mit der imaginären Achse zusammenfällt und $z_0 < 0$, $s = 0$ sowie $\operatorname{Re} W(T) \geqq 0$ ist. Dann erhält man für eine reelle Zahl $\gamma > 0$ und für $\|x\| = 1$ die Ungleichung

$$\|Tx - \gamma x\|^2 = \|Tx\|^2 + \gamma^2 \|x\|^2 - 2\gamma \cdot \operatorname{Re}(Tx, x) \leqq \|T\|^2 + \gamma^2,$$

somit

$$\|T - \gamma I\|^2 \leqq \|T\|^2 + \gamma^2.$$

Nach einem bekannten Satz von v. NEUMANN (vgl. [3] oder [4]) ist dann die Kreis-

[1]) Mit ∂M bezeichnen wir den Rand einer Menge M.

scheibe $C_{\gamma,r(\gamma)}$ mit dem Mittelpunkt γ und dem Radius $r(\gamma) = (\|T\|^2+\gamma^2)^{1/2}$ eine Spektralmenge von T. Andererseits gilt

$$r(\gamma)-\gamma = \gamma\cdot\left[\left(1+\frac{\|T\|^2}{\gamma^2}\right)^{1/2}-1\right] = \gamma\cdot O\left(\frac{1}{\gamma^2}\right)\to 0 \quad \text{mit} \quad \gamma\to\infty.$$

Für hinreichend großes γ ist also z_0 nicht in $C_{\gamma,r(\gamma)}$ enthalten. Damit ist der Satz bewiesen.

Im Zusammenhang mit Satz 1 ist vielleicht das folgende Resultat überraschend, das in [2] bewiesen wird:

Satz 2. *Falls für eine Folge $\{z_n\}$ von Extremalpunkten z_n der konvexen Menge $\overline{W(T)}$, die in der Menge aller Extremalpunkte von $\overline{W(T)}$ dicht liegen, eine Folge von Kreisscheiben C_n existiert, so daß $z_n\in\partial C_n$ und C_n Spektralmenge von T für alle $n=1,2,\dots$ ist, so ist $\overline{W(T)}$ gleich der konvexen Hülle des Spektrums von T.*

Bezeichne nun $\mathfrak{N}_1(T)$ die Menge aller starken normalen Dilatationen N des Operators T von der Form $N=\alpha U+\beta I$, wo U ein unitärer Operator, I die Identität und α,β komplexe Zahlen sind.

Satz 3. *Für jeden beschränkten linearen Operator T in $\mathfrak{H}$ gilt*

$$W(T)\subseteqq \bigcap_{N\in\mathfrak{N}_1(T)} W(N)\subseteqq \overline{W(T)}.$$

Beweis. Für eine Dilatation N von T und einen Vektor $x\in\mathfrak{H}$ ist $(Tx,x)=(Nx,x)$, somit $W(T)\subseteqq W(N)$. Daher brauchen wir nur $\bigcap_{N\in\mathfrak{N}_1(T)} W(N)\subseteqq\overline{W(T)}$ zu beweisen. Falls die Kreisscheibe $C_{\gamma,r}=\{z:|z-\gamma|\leqq r\}$ Spektralmenge von T ist, so ist $\|r^{-1}(T-\gamma I)\|\leqq 1$, folglich gibt es nach [5] eine starke unitäre Dilatation U von $r^{-1}(T-\gamma I)$, woraus wegen $T^n=\sum_0^n\binom{n}{k}\gamma^{n-k}(T-\gamma I)^k$ sofort folgt, daß $N=rU+\gamma I$ eine starke normale Dilatation von T ist. Ferner ist für einen normalen Operator der Abschluß des numerischen Wertebereichs gerade die konvexe Hülle seines Spektrums, also $\overline{W(N)}\subseteqq C_{\gamma,r}$. Hieraus ergibt sich unter Berücksichtigung von Satz 1 die Behauptung.

3. Schließlich bezeichnen wir mit $\mathfrak{U}(T)$ die Menge der unitären Dilatationen eines Operators T.

Satz 4. *Wenn sich $\overline{W(T)}$ als Durchschnitt von konvexen Spektralmengen des Operators T darstellen läßt, die sämtlich im Einheitskreis $\{z:|z|\leqq 1\}$ liegen, so gilt*

$$(*)\qquad W(T)\subseteqq \bigcap_{U\in\mathfrak{U}(T)} W(U)\subseteqq\overline{W(T)}.$$

Beweis. Aus demselben Grunde wie in Satz 3 genügt es, $\bigcap_{U\in\mathfrak{U}(T)} W(U)\subseteqq\overline{W(T)}$ zu zeigen. Sei also $z_0\notin\overline{W(T)}$ und $|z_0|\leqq 1$. Dann gibt es eine konvexe und kompakte, ganz in der Einheitskreisscheibe $C_{0,1}=\{z:|z|\leqq 1\}$ gelegene Spektralmenge X von T, die z_0 nicht enthält. Eine geeignete Stützgerade g an X trennt dann X von z_0. Die Gerade g teilt $C_{0,1}$ in zwei Teile. Wir betrachten davon den (abgeschlossenen)

Teil S, der X enthält, aber nicht z_0. Der Rand ∂S ist von der Form $\partial S = J \cup B$, wobei J eine Strecke auf g und B ein Bogen auf dem Einheitskreis $\partial C_{0,1}$ ist. Mit X ist auch S eine Spektralmenge von T. Auf Grund des Satzes 6 in Sz.-Nagy—Foiaş [6] gibt es eine (starke) normale Dilatation N von T mit $\sigma(N) \subseteqq \partial S$. Nach dem Spektraltheorem kann man N in eine orthogonale Summe $N = N_1 \oplus U_2$ von normalen Operatoren N_1 und U_2 mit

$$\text{(i)}\quad \sigma(N_1) \subseteqq J, \qquad \text{(ii)}\quad \sigma(U_2) \subseteqq B$$

zerlegen. Aus (ii) folgt sofort, daß U_2 unitär ist, und aus (i) ergibt sich, daß für eine lineare Abbildung $p(z) = \alpha z + \beta$, die J auf das Intervall $\{\lambda : 0 \leqq \lambda \leqq 1\}$ abbildet, der durch $A_1 = p(N_1)$ definierte Operator selbstadjungiert und $0 \leqq A_1 \leqq I$ ist. Mittels einer einfachen Konstruktion (vgl. [4]) findet man einen Projektor P_1 als Dilatation von A_1. Das Spektrum eines Projektors besteht höchstens aus den Punkten 0 und 1. Der Operator N_1 besitzt daher die normale Dilatation $U_1 = q(A_1)$, $\left(q(z) = \frac{z-\beta}{\alpha}\right)$, deren Spektrum höchstens aus den Endpunkten der Strecke J besteht, also auf dem Rande des Einheitskreises liegt. Dann ist U_1 aber unitäre Dilatation von N_1 und hierauf $U = U_1 \oplus U_2$ unitäre Dilatation von $N = N_1 \oplus U_2$ und damit erst recht von T, und $\sigma(U) = \sigma(U_1) \cup \sigma(U_2) \subseteqq B \subseteqq \partial S$. Hieraus folgt $\overline{W(U)} \doteq$ konvexe Hülle von $\sigma(U) \subseteqq S$. Damit haben wir für den Punkt $z_0 \notin \overline{W(T)}$ mit $|z_0| \leqq 1$ eine unitäre Dilatation U von T gefunden, so daß $z_0 \notin \overline{W(U)}$ ist. Da der numerische Wertebereich eines unitären Operators im Einheitskreis $C_{0,1}$ liegt, erhält man hieraus $\bigcap_{U \in \mathfrak{U}(T)} W(U) \subseteqq \overline{W(T)}$, womit alles gezeigt ist.

Zusatz bei der Korrektur: Dr. C. Berger hat mir inzwischen mitgeteilt, daß er die Richtigkeit von ($*$) für alle Kontraktionen T bewiesen hat. Ferner hat er — unabhängig von mir und noch nicht veröffentlicht — ebenfalls die Ergebnisse von Satz 1, 3 gefunden.

Literaturverzeichnis

[1] P. R. Halmos, Numerical ranges and normal dilations, *Acta Sci. Math.*, **25** (1964), 1—5.

[2] S. Hildebrandt, *Numerischer Wertebereich und Spektralmengen* (zu erscheinen).

[3] J. von Neumann, Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes, *Math. Nachrichten*, **4** (1951), 258—281.

[4] F. Riesz—B. Sz.-Nagy, *Vorlesungen über Funktionalanalysis* (Berlin, 1956).

[5] B. Sz.-Nagy, Sur les contractions de l'espace de Hilbert, *Acta Sci. Math.*, **15** (1953), 87—92.

[6] B. Sz.-Nagy—C. Foiaş, Sur les contractions de l'espace de Hilbert. III, *Acta Sci. Math.*, **19** (1958), 26—45.

(Eingegangen am 12. Oktober 1964)

Positive definite kernels generated by operator-valued analytic funtions

By BÉLA SZ.-NAGY in Szeged

Let $\Theta(z)$ be a function of the complex variable z in the unit disc $D_0 = \{z: |z|<1\}$, whose values are linear transformations of a Hilbert space $\mathfrak{H}$ into a Hilbert space $\mathfrak{H}'$, of norm $\|\Theta(z)\| \leqq 1$, and which is analytic in D_0:

$$\Theta(z) = A_0 + zA_1 + z^2A_2 + \dots \quad (z \in D_0).$$

This function generates the "kernel"

$$K(u, v) = \frac{I - \Theta(v)^* \Theta(u)}{1 - \bar{v}u} \qquad (u, v \in D_0). \tag{1}$$

In his Yale dissertation 1963 ROVNYAK has proved, among other things, the following proposition (*cf.* [1] theorem 4. 3):

The kernel $K(u, v)$ is positive definite, i. e.

$$\sum_m \sum_n (K(u_m, u_n)h_m, h_n) \geqq 0 \tag{2}$$

for every finite system $\{u_n\}_1^N$ of points of D_0, and every corresponding system $\{h_n\}_1^N$ of vectors in $\mathfrak{H}$.

The proof given by ROVNYAK is rather indirect, so it might be of some interest to present here a simple proof.

This proof is based on the following *formula:*

$$(K(u, v)h, k) = \tag{3}$$
$$= \frac{r^2 - \bar{v}u}{1 - \bar{v}u} \frac{1}{2\pi} \int_0^{2\pi} \left[\left(\frac{\Theta(z) - \Theta(u)}{z - u} h, \frac{\Theta(z) - \Theta(v)}{z - v} k \right) + \left(\frac{\Omega(z)}{z - u} h, \frac{1}{z - v} k \right) \right] dt$$

where $z = re^{it}$; $\max\{|u|, |v|\} < r < 1$; $h, k \in \mathfrak{H}$; and

$$\Omega(z) = I - \Theta(z)^* \Theta(z) \quad (\Omega(z) \geqq O).$$

Indeed, if we choose r such that $\max\{|u_n|\}_1^N < r < 1$, we shall have, in virtue of (3),

$$\sum_m \sum_n (K(u_m, u_n)h_m, h_n) =$$
$$= \lim_{r \to 1} \frac{1}{2\pi} \int_0^{2\pi} \left[\left\| \sum_n \frac{\Theta(z) - \Theta(u_n)}{z - u_n} h_n \right\|^2 + \left(\Omega(z) \sum_n \frac{1}{z - u_n} h_n, \sum_n \frac{1}{z - u_n} h_n \right) \right] dt$$

($z = re^{it}$); since $\Omega(z) \geqq O$, this yields (2).

Now, to prove (3), we first observe that

$$\left(\frac{\Theta(z)}{z-u}h, \frac{\Theta(z)}{z-v}k\right)+\left(\frac{\Omega(z)}{z-u}h, \frac{1}{z-v}k\right)=\frac{1}{z-u}\frac{1}{\bar{z}-\bar{v}}(h, k),$$

and

$$\left(\frac{\Theta(u)}{z-u}h, \frac{\Theta(v)}{z-v}k\right)=\frac{1}{z-u}\frac{1}{\bar{z}-\bar{v}}(\Theta(u)h, \Theta(v)k),$$

$$\frac{1}{2\pi}\int_0^{2\pi}\frac{dt}{(z-u)(\bar{z}-\bar{v})}=\frac{1}{r^2}\cdot\frac{1}{2\pi i}\oint_{|\zeta|=1}\frac{d\zeta}{\left(\zeta-\frac{u}{r}\right)\left(1-\frac{\bar{v}}{r}\zeta\right)}=\frac{1}{r^2}\frac{1}{1-\frac{\bar{v}}{r}\frac{u}{r}}=\frac{1}{r^2-\bar{v}u}.$$

Next we obtain

$$\frac{1}{2\pi}\int_0^{2\pi}\left(\frac{\Theta(z)}{z-u}h, \frac{\Theta(v)}{z-v}k\right)dt=\frac{1}{r^2}\frac{1}{2\pi i}\oint_{|\zeta|=1}\frac{(\Theta(r\zeta)h, \Theta(v)k)}{\left(\zeta-\frac{u}{r}\right)\left(1-\frac{\bar{v}}{r}\zeta\right)}d\zeta=$$

$$=\frac{1}{r^2}\frac{\left(\Theta\left(r\frac{u}{r}\right)h, \Theta(v)k\right)}{1-\frac{\bar{v}}{r}\frac{u}{r}}=\frac{(\Theta(u)h, \Theta(v)k)}{r^2-\bar{v}u}$$

and analogously, passing through complex conjugates,

$$\frac{1}{2\pi}\int_0^{2\pi}\left(\frac{\Theta(u)}{z-u}h, \frac{\Theta(z)}{z-v}k\right)dt=\frac{(\Theta(u)h, \Theta(v)k)}{r^2-\bar{v}u}.$$

Putting these results together we obtain the desired result (3).

Remark. In the case the space $\mathfrak{H}$ is separable, the radial limit $\Theta(e^{it})=$ $=\lim\limits_{r\to 1}\Theta(re^{it})$ exists almost everywhere, in the sense of strong operator convergence, *cf.* [2], n° 1. Hence $\Omega(e^{it})=\lim\limits_{r\to 1}\Omega(re^{it})$ also exists a. e., at least in the sense of weak operator convergence. In virtue of LEBESGUE's theorem on the integration of bounded sequences of functions we obtain from (3) as $r\to 1$:

$$(K(u, v)h, k)=$$
$$=\frac{1}{2\pi}\int_0^{2\pi}\left[\left(\frac{\Theta(e^{it})-\Theta(u)}{e^{it}-u}h, \frac{\Theta(e^{it})-\Theta(v)}{e^{it}-v}k\right)+\left(\frac{\Omega(e^{it})}{e^{it}-u}h, \frac{1}{e^{it}-v}k\right)\right]dt.$$

(Added by proof reading.) An alternative simple proof of (2) was kindly indicated to me in a letter of February 10, 1965, by DR. ROVNYAK.

References

[1] J. ROVNYAK, Some Hilbert spaces of analytic functions, *Yale dissertation* (1963).
[2] B. SZ.-NAGY—C. FOIAŞ, Sur les contractions de l'espace de Hilbert. VIII. Fonctions caractéristiques. Modèles fonctionnels, *Acta Sci. Math.*, **25** (1964), 38—71.

(Received December 1, 1964)

Corrections et compléments aux Contractions IX

Par BÉLA SZ.-NAGY à Szeged et CIPRIAN FOIAŞ à Bucarest

1. Dans la Note [IX], Proposition 5. 4 est fausse. Un contre-exemple est fourni par la fonction analytique contractive extérieure $\{E^2, E^1, \Theta(\lambda)\}$ où

$$\Theta(\lambda)\begin{pmatrix}x_1\\x_2\end{pmatrix}=\left(\frac{1}{2\sqrt{2}}(\lambda-1), \frac{1}{\sqrt{2}}\lambda\right)\begin{pmatrix}x_1\\x_2\end{pmatrix} \qquad (cf.\ [\mathrm{IX}],\ \mathrm{p.}\ 313);$$

pour le vecteur $f=\begin{pmatrix}0\\1\end{pmatrix}$ on a notamment $\Theta(0)f=0$ sans que $\Theta(\lambda)f$ s'annule identiquement. (On a commis l'erreur lorsqu'on a fait usage, p. 308, de la linéarité de p_f en fonction de f, ce qui, en général, n'est pas le cas.)

Parmi les propositions et théorèmes qui suivent dans [IX], Théorème 3 (sur l'existence d'une variété de sous-espaces invariants pour une certaine classe d'opérateurs), ainsi que sa démonstration, ne dépendent pas de Proposition 5. 4, donc ils ne sont pas touchés par cette erreur. Par contre, Propositions 5. 5 et 5. 6 et Théorèmes 2 et 4 dépendent partiellement de Proposition 5. 4, donc ils doivent être réexaminés. En laissant ouvert le problème de décider s'ils restent valables dans toute leur généralité ou non, nous nous contentons ici de montrer qu'ils subsistent dans le cas d'indices de défaut finis.

2. Observons d'abord que l'énoncé suivant subsiste avec sa démonstration indiquée dans [IX], p. 310—311:

Proposition A. *Pour toute fonction contractive extérieure* $\{\mathfrak{E}, \mathfrak{E}_*, \Theta(\lambda)\}$ *on a*

(i) $\overline{\Theta(\lambda)\mathfrak{E}}=\mathfrak{E}_*$ *pour tout* $\lambda \in D_0$, [1])

(ii) $\overline{\Theta(e^{it})\mathfrak{E}}=\mathfrak{E}_*$ *pour presque tous les points* $t \in (0, 2\pi)$.

Dans le cas où $\dim \mathfrak{E}_* < \infty$, (i) et (ii) veulent dire le même que $\Theta(\lambda)\mathfrak{E}=\mathfrak{E}_*$, $\Theta(e^{it})\mathfrak{E}=\mathfrak{E}_*$. Il s'ensuit que si $\dim \mathfrak{E}_* < \infty$ et $\dim \mathfrak{E} \neq \dim \mathfrak{E}_*$, $\Theta(\lambda)$ ne peut avoir d'inverse pour aucun $\lambda \in D_0$, même pas au sens large.

Pour le cas $\dim \mathfrak{E} = \dim \mathfrak{E}_* = n < \infty$ on a la

Proposition B. *Pour que la fonction analytique contractive matricielle* $\Theta(\lambda)$ *de type* $n \times n$ *soit extérieure, il faut et il suffit que la fonction scalaire bornée* $d(\lambda) = \det \Theta(\lambda)$ *soit extérieure.* (*Cf.* [1], Lecture XI.)

[1]) D_0 est le disque unité ouvert dans le plan des nombres complexes.

Bien entendu, la condition que $d(\lambda)$ soit extérieure veut dire que $\overline{dH^2(E^1)} = H^2(E^1)$; d'après le théorème de BEURLING cela est équivalent à la condition que $\log |d(e^{it})|$ soit intégrable et que la relation suivante soit vérifiée:

$$(1) \qquad d(\lambda) = \varkappa \exp\left[\frac{1}{2\pi}\int_0^{2\pi} \frac{e^{it}+\lambda}{e^{it}-\lambda} \log |d(e^{it})|\, dt\right], \qquad |\varkappa| = 1.$$

On a

$$(2) \qquad \Theta(\lambda)^{-1} = \frac{1}{d(\lambda)}\, \Theta^{A}(\lambda)$$

où $\Theta^A(\lambda)$ est la matrice algébriquement adjointe à $\Theta(\lambda)$. Puisque $\|\Theta(\lambda)\| \leqq 1$, les éléments de $\Theta(\lambda)$ sont bornés dans D_0, d'où il s'ensuit que $\Theta^A(\lambda)$ est aussi bornée dans D_0.

Comme pour $\Theta^{\sim}(\lambda) = \Theta(\bar{\lambda})^*$ le déterminant $d^{\sim}(\lambda)$ est le conjugué complexe de $d(\bar{\lambda})$, $d^{\sim}(\lambda)$ est une fonction extérieure en même temps que $d(\lambda)$. Par conséquent, $\Theta^{\sim}(\lambda)$ est extérieure en même temps que $\Theta(\lambda)$.

Ces résultats établissent la Proposition 5. 5 de [IX] dans le cas où $\mathfrak{E}_*$ est de dimension finie et cela dans la forme précisée suivante:

Proposition 5. 5*. *Soit* $\{\mathfrak{E}, \mathfrak{E}_*, \Theta(\lambda)\}$ *une fonction analytique contractive extérieure avec* $\dim \mathfrak{E}_* < \infty$. *Il y a alors deux possibilités: ou bien* $\Theta(\lambda)^{-1}$ *existe au sens strict en chaque point de* D_0, *ou bien* $\Theta(\lambda)^{-1}$ *n'existe pour aucun point de* D_0, *même pas au sens large. Le premier cas se présente lorsque* $\dim \mathfrak{E} = \dim \mathfrak{E}_*$, *et le second lorsque* $\dim \mathfrak{E} \neq \dim \mathfrak{E}_*$. *Dans le premier cas la fonction* $\Theta^{\sim}(\lambda)$ *est aussi extérieure.*

On obtient alors Théorème 2 de [IX] dans la forme restreinte suivante:

Théorème 2*. *Soit* $T \in C_{\cdot 1}$, *avec les indices de défaut* $\mathfrak{d}_T, \mathfrak{d}_{T^*}$ [2]) *dont* $\mathfrak{d}_{T^*}$ *est fini. Il y a deux possibilités: ou bien chaque point de* D_0 *est une valeur propre de* T, *ou bien aucun point de* D_0 *n'appartient au spectre de* T: *le premier cas se présente lorsque* $\mathfrak{d}_T \neq \mathfrak{d}_{T^*}$, *et le second lorsque* $\mathfrak{d}_T = \mathfrak{d}_{T^*}$. *Dans le premier cas* $T \notin C_{11}$ *et dans le second cas* $T \in C_{11}$.

Démonstration. En vertu de [VIII], Corollaire, p. 58, $T \in C_{\cdot 1}$ veut dire que la fonction caractéristique $\{\mathfrak{D}_T, \mathfrak{D}_{T^*}, \Theta_T(\lambda)\}$ de T est une fonction extérieure; de plus on a $\dim \mathfrak{D}_{T^*} = \mathfrak{d}_{T^*} < \infty$. Le théorème résulte alors, sauf la dernière assertion, de la Proposition 5. 5* et du Théorème 4 de [VIII]. Quant à la dernière assertion, observons que dans le second cas $\Theta_T^{\sim}(\lambda)$ est, toujours en vertu de la Proposition 5. 5*, aussi extérieure, donc $T \in C_{1\cdot}$ et par conséquent $T \in C_{11}$. Dans le premier cas tout point $\lambda \in D_0$ étant une valeur propre, il y a des vecteurs $x_\lambda \neq 0$ tels que $Tx_\lambda = \lambda x_\lambda$, d'où $T^n x_\lambda = \lambda^n x_\lambda \to 0$ $(n \to \infty)$, ce qui exclut que T appartienne à C_{11}. Cela achève la démonstration.

---

[2]) $\mathfrak{d}_T = \dim \mathfrak{D}_T$ où $\mathfrak{D}_T = \overline{(I - T^*T)^{1/2}\mathfrak{H}}$, et analoguement pour T^*.

3. Proposition 5. 6 de [IX] est valable dans le cas où dim $\mathfrak{E}_* < \infty$:

Proposition 5. 6*. *Soit $\{\mathfrak{E}, \mathfrak{E}_*, \Theta(\lambda)\}$ une fonction contractive extérieure telle que* dim $\mathfrak{E}_* < \infty$ *et que $\Theta(e^{it})$ est isométrique pour presque tous les points e^{it} d'un arc ω de la circonférence unité. $\Theta(\lambda)$ se prolonge alors analytiquement au travers de ω à tout l'extérieur du cercle unité.*

Démonstration. De Proposition A (ii) il s'ensuit que $\Theta(e^{it})$ est même unitaire pp. sur ω et que, par conséquent, dim $\mathfrak{E}$ = dim $\mathfrak{E}_* < \infty$. Pour $d(\lambda) = \det \Theta(\lambda)$ on a alors $|d(e^{it})| = 1$ pp. sur ω; en vertu de la formule (1) cela entraîne que

$$\log |d(\lambda)| = \frac{1}{2\pi} \int' P(r, \tau - t) \log |d(e^{it})| \, dt$$

où $\lambda = re^{i\tau}$ $(0 \leqq r < 1)$, $P(r, t)$ est le noyau de Poisson, et l'intégrale est prise sur l'ensemble des points t tels que $e^{it} \in \omega'$ où ω' est l'arc complémentaire à ω. De là il s'ensuit que la fonction $1/d(\lambda)$ est bornée dans tout sous-ensemble de D_0, qui est à distance positive à ω'; en vertu de (2) il en est alors de même pour $\Theta(\lambda)^{-1}$. La démonstration s'achève comme indiquée aux p. 311—312 de [IX], faisant usage du principe de réflexion de Schwarz.

Théorème 4*. *Soit $T \in C_{\cdot 1}$ telle que l'indice de défaut $\mathfrak{d}_{T^*}$ est fini et que $\sigma(T)$ ne recouvre pas la corconférence unité. On a alors $\sigma(T) = \sigma(U^\circ)$.*

(U° est la partie "résiduelle" de la dilatation unitaire minimum de T, envisagée dans Théorème 3.) La démonstration se fait par la méthode indiquée dans [IX], p. 313, faisant usage de la Proposition 5. 6*.

4. Finalement, le Corollaire à la p. 315 de [IX] découle, dans le cas où $T \in C_{11}$ a ses deux indices de défaut finis, des Théorèmes 3 et 4* de la manière indiquée l. c., si l'on fait usage encore du fait que les opérateurs T_β y envisagés ont aussi leurs indices de défaut finis. Cela résulte de la proposition suivante qui présente un complément au n° 4 de [IX], intéressant en soi-même:

Proposition C. *Soit T une contraction complètement non-unitaire de $\mathfrak{H}$, aux indices de défaut (finis ou non) $\mathfrak{d}_T$ et $\mathfrak{d}_{T^*}$. Soit $\mathfrak{H}_1$ $(\neq \{0\})$ un sous-espace de $\mathfrak{H}$, invariant pour T. L'opérateur $T_1 = T|\mathfrak{H}_1$ a alors ses indices de défaut $\mathfrak{d}_{T_1}$, $\mathfrak{d}_{T_1^*}$ tels que*

$$\mathfrak{d}_{T_1} \leqq \mathfrak{d}_T \text{ et } \mathfrak{d}_{T_1^*} \leqq \mathfrak{d}_T + \mathfrak{d}_{T^*}.$$

Démonstration. Soit $\{\mathfrak{E}, \mathfrak{E}_*, \Theta(\lambda)\}$ une fonction analytique contractive qui coïncide avec la fonction caractéristique de T et soit $\Theta(\lambda) = \Theta_2(\lambda)\Theta_1(\lambda)$ la factorisation de cette fonction en produit de deux fonctions analytiques contractives, $\{\mathfrak{E}, \mathfrak{F}, \Theta_1(\lambda)\}$ et $\{\mathfrak{F}, \mathfrak{E}_*, \Theta_2(\lambda)\}$, correspondant au sous-espace invariant $\mathfrak{H}_1$ au sens du Théorème 1 de [IX]. On sait, *cf.* [IX], p. 295 et Proposition 4. 4, que l'application

$$\Delta(t)g \to \Delta_2(t)\Theta_1(e^{it})g \oplus \Delta_1(t)g \qquad (g \in \mathfrak{E})$$

est isométrique et se prolonge par continuité à une application unitaire de $\overline{\Delta(t)\mathfrak{E}}$

sur $\overline{\Delta_2(t)\mathfrak{F}} \oplus \overline{\Delta_1(t)\mathfrak{E}}$, pour presque tous les t. Cela entraîne $\dim \overline{\Delta(t)\mathfrak{E}} = \dim \overline{\Delta_2(t)\mathfrak{F}} + \dim \overline{\Delta_1(t)\mathfrak{E}}$ pp. et par conséquent

(3) $$\dim \overline{\Delta_2(t)\mathfrak{F}} \leqq \dim \overline{\Delta(t)\mathfrak{E}} \leqq \dim \mathfrak{E} = \mathfrak{d}_T \quad \text{pp.}$$

D'autre part, de la relation

$$I_{\mathfrak{F}} = \Delta_2(t)^2 + \Theta_2(e^{it})^* \Theta_2(e^{it}) \quad \text{pp.}$$

il s'ensuit que, pour presque tous les t, $\mathfrak{F}$ est sous-tendu par $\Delta_2(t)\mathfrak{F}$ et $\Theta_2(e^{it})^* \Theta_2(e^{it})\mathfrak{F}$, donc

(4) $$\dim \mathfrak{F} \leqq \dim \overline{\Delta_2(t)\mathfrak{F}} + \dim \overline{\Theta_2(e^{it})^* \Theta_2(e^{it})\mathfrak{F}}.$$

Comme $\Theta_2(e^{it})\mathfrak{F} \subset \mathfrak{E}_*$, on a $\dim \overline{\Theta_2(e^{it})\mathfrak{F}} \leqq \dim \mathfrak{E}_* = \mathfrak{d}_{T^*}$; par conséquent

(5) $$\dim \overline{\Theta_2(e^{it})^* \Theta_2(e^{it})\mathfrak{F}} \leqq \dim \overline{\Theta_2(e^{it})\mathfrak{F}} \leqq \mathfrak{d}_{T^*}.$$

Ainsi, de (3), (4) et (5) il résulte:

(6) $$\dim \mathfrak{F} \leqq \mathfrak{d}_T + \mathfrak{d}_{T^*}.$$

Or, on sait (*cf.* Proposition 4. 3 de [IX]) que la fonction caractéristique de T_1 coïncide avec la partie pure $\{\mathfrak{E}^\circ, \mathfrak{F}^\circ, \Theta_1^\circ(\lambda)\}$ de $\{\mathfrak{E}, \mathfrak{F}, \Theta_1(\lambda)\}$ où $\mathfrak{E}^\circ \subset \mathfrak{E}$, $\mathfrak{F}^\circ \subset \mathfrak{F}$. De là il résulte que

$$\mathfrak{d}_{T_1} = \dim \mathfrak{E}^\circ \leqq \dim \mathfrak{E} = \mathfrak{d}_T,$$

$$\mathfrak{d}_{T_1^*} = \dim \mathfrak{F}^\circ \leqq \dim \mathfrak{F} \leqq \mathfrak{d}_T + \mathfrak{d}_{T^*},$$

ce qui achève la démonstration de Proposition C.

Ouvrages cités

[1] H. Helson, *Lectures on invariant subspaces* (New York—London, 1964).
[VIII] B. Sz.-Nagy—C. Foiaş, Sur les contractions de l'espace de Hilbert. VIII. Fonctions caractéristiques. Modèles fonctionnels, *Acta Sci. Math.*, **25** (1964), 38—71.
[IX] B. Sz.-Nagy—C. Foiaş, Sur les contractions de l'espace de Hilbert. IX. Factorisations de la fonction caractéristique. Sous-espaces invariants, *Acta Sci. Math.*, **25** (1964), 283—316).

(Reçu le 10 février 1965)

Bibliographie

Demetrios A. Kappos, Strukturtheorie der Wahrscheinlichkeitsfelder und -Räume (Ergebnisse der Mathematik und ihrer Grenzgebiete, Neue Folge, Heft 24), 136 Seiten, Berlin—Göttingen—Heidelberg, Springer-Verlag, 1960.

Die Idee, die Wahrscheinlichkeit als ein Maß auf einer Booleschen Algebra zu definieren, stammt ursprünglich von W. I. GLIVENKO[1]). Zur selben Zeit hat C. CARATHÉODORY[2]) eine Maß- und Integraltheorie auf Booleschen Algebren entwickelt. Diese Richtung wurde seitdem wesentlich weiterentwickelt, hauptsächlich durch den Verfasser des vorliegenden Buches, ferner von A. N. KOLMOGOROFF[3]) und von anderen. Es ist wohlbekannt, daß diese Auffassung gewisse Vorteile hat: zum Beispiel gibt es in dieser Auffassung keinen Unterschied zwischen dem unmöglichen Ereignis und einem Ereignis mit der Wahrscheinlichkeit Null.

Das Buch besteht aus 8 Kapiteln. Im I. Kapitel wird der Begriff eines Wahrscheinlichkeitsfeldes (kurz: w-Feldes) eingeführt. Ein w-Feld ist ein Boolescher Ring F mit Einheit, auf welchem eine strikt positive, normierte und additive Funktion w, die Wahrscheinlichkeit definiert ist. Falls von w statt strikter Positivität nur Nichtnegativität verlangt wird, spricht man von einem Quasi-w-Feld. Es folgen einige Beispiele und Existenzsätze; das Kapitel endet mit dem Stoneschen Darstellungssatz.

Kapitel II beschäftigt sich mit unendlichen Operationen in w-Feldern. Die w-Felder (F, w), wo F ein σ-Boolescher Ring und w σ-additiv ist, werden σ-w-Felder gennant. Die Frage der Erweiterung eines w-Feldes zu einem σ-w-Feld wird gründlich diskutiert.

Kapitel III beschäftigt sich mit dem Verhältnis der in den ersten zwei Kapiteln entwickelten Theorie der Wahrscheinlichkeitsfelder zu der Kolmogoroffschen Wahrscheinlichkeitstheorie. Sind Ω eine nichtleere Menge, K ein Boolescher Ring, der aus Teilmengen von Ω besteht und selbst Ω als Element enthält, und v eine stetige Quasi-Wahrscheinlichkeit auf K, so wird $[\Omega, K, v]$ ein Wahrscheinlichkeitsraum (kurz: w-Raum) genannt.

Es wird gezeigt, daß jedes w-Feld durch einen w-Raum dargesellt werden kann, derart, daß die Wahrscheinlichkeit bei dieser Darstellung die Eigenschaft der σ-Additivität besitzt, weiter, daß umgekehrt man aus einem w-Raum durch Bildung von Restklassen modulo Nullmengen in ein w-Feld übergehen kann. Die beiden Theorien sind also äquivalent.

Kapitel IV bis VI beschäftigen sich mit dem Begriff des Cartesischen Produkts von w-Feldern, und mit dem Begriff der Unabhängigkeit.

Kapitel VII beschäftigt sich mit topologischen bzw. kompakten w-Räumen.

Kapitel VIII beschäftigt sich mit dem Begriff der bedingten Wahrscheinlichkeitsräumen[4]) (kurz: bedingte w-Räume). Es sei Ω eine nichtleere Menge, K ein σ-Körper von Teilmengen von Ω, der auch Ω als Element enthält, und T ein nichtleeres Untersystem von K. Auf dem Cartesischen Produkt $K \times T$ sei eine nichtnegative Funktion $P(A|B)$ $(A \in K, B \in T)$ definiert, für die $P(B|B)=1$ für $B \in T$ gilt und die bei festem B als Funktion von A ein σ-additives Maß auf K ist, und für welche

---

[1]) W. I. GLIVENKO, Théorie générale des structures, *Actualités Sci. et Industr.*, Nr. 652 (Paris, 1938).

[2]) C. CARATHÉODORY, Entwurf für eine Algebraisierung des Integralbegriffes, *Sitzungsber. Math.-Naturw. Klasse Bayer. Akad. Wiss. München*, **1938**, 24—28.

[3]) A. KOLMOGOROFF, Algèbres de Boole métriques complètes, *VI. Zjazd Matematyków Polskich, Kraków*, **1950**, 22—30.

[4]) Siehe A. RÉNYI, On a new axiomatic theory of probability, *Acta Math. Acad. Sci. Hung.*, **6** (1955), 285—335 und A. RÉNYI, On conditional probability spaces generated by a dimensionally ordered set of measures, *Teorija Verojatn. Prim.*, **1** (1956), 61—71.

aus $A \in K$, $B \in K$, $C \in T$ und $BC \in T$ folgt: $P(A|BC)P(B|C)=P(AB|C)$. Dann wird $[\Omega, K, T, P]$ ein bedingter w-Raum und $P(A|B)$ die bedingte Wahrscheinlichkeit von A unter der Bedingung B genannt. Die Frage der Erweiterung eines bedingten w-Raumes wird diskutiert, es werden ferner diejenigen bedingten w-Räume charakterisiert, welche vom einfachen Quotiententypus sind, d. h. für welche ein σ-additives Maß μ auf K gibt, derart, daß $\mu(B)>0$ für $B \in T$ gilt und

$$P(A|B) = \frac{\mu(AB)}{\mu(B)}$$

für jede $A \in K$, $B \in T$ ist. Ferner werden die Resultate von Á. Császár[5]) (dessen Name an mehreren Stellen falsch gesetzt wurde) über die Erzeugung von allgemeinen bedingten w-Feldern durch eine Familie von σ-additiven Maßen dargestellt.

Das Buch enthält einen Anhang, in welchem die nötigen algebraischen Vorkenntnisse zusammengestellt sind, ferner ein Literaturverzeichnis von 65 Arbeiten, und ein Namen- und Sachverzeichnis.

Das sehr klar und sorgfältig geschriebene, wertvolle und nützliche Buch, in welchem alle das Thema betreffenden Ergebnisse zum ersten Male gesammelt, systematisch dargestellt, miteinander verglichen und in vieler Hinsicht ergänzt sind, kann jedem, der sich für die Grundlagen der Wahrscheinlichkeitstheorie interessiert, warm empfohlen werden.

A. Rényi (Budapest)

G. Scorza Dragoni, Elemente di analisi matematica. Vol. I. **Elementi di algebra,** VIII+584 p.; Vol. II. **La continuità e la differenziabilità,** VI+692 p.; Vol. III. **La teoria elementare dell'integrazione,** VI+584 p., Padova, CEDAM, 1961—62. — Lit. 5000+6000+5000.

Das dreibändige ausgezeichnete Lehrbuch des Verfassers umfaßt das übliche Material der einleitenden Algebra- und Analysisvorlesungen an den Universitäten. Das Lehrbuch verdient deshalb eine besondere Aufmerksamkeit, da es die strengsten Anforderungen der mathematischen Exaktheit immer vor Augen hält, und diese Exaktheit mit einer klaren, pädagogisch durchdachten Vortragsmethode verbindet. Obwohl das behandelte Material als klassisch angesehen werden kann, ein anderer großer Vorteil des Buches ist, daß es die eingeführten Begriffe — den Möglichkeiten angemessen — von dem Standpunkt der gegenwärtigen Lage der Wissenschaft behandelt.

Der erste Band beginnt mit der Untersuchung der Mengen, dann baut er die reellen und komplexen Zahlenkörper axiomatisch auf. Danach folgen die Grundbegriffe der Kombinatorik, der Determinantentheorie und der Matrizenrechnung. Es kommen weiterhin die Grundlagen der Gruppen-, Ring- und Körpertheorie, die lineare Unabhängigkeit und Abhängigkeit vor, dann folgt die Behandlung der Lösbarkeit der linearen Gleichungssysteme. Es wird ein Kapitel den numerischen linearen Räumen, vier Kapitel den Euklidischen Räumen gewidmet. Ein Kapitel befaßt sich mit über dem reellen bzw. komplexen Zahlenkörper definierten ein- und mehrveränderlichen Polynomen, dann folgt eine didaktisch wertvolle Beweisführung des Grundsatzes der Algebra in mehreren Schritten. Dem folgt die Untersuchung der über dem rationalen bzw. reellen Zahlenkörper definierten algebraischen Gleichungen. Der erste Band schließt mit der Einführung in die Theorie der quadratischen Formen.

Der zweite Band beginnt mit der Behandlung der Punktmengen in den Euklidischen Räumen. Dem folgen die Begriffe und Sätze der Grenzwerte in der Reihenfolge: Zahlenfolge, Zahlenreihen, reelle und komplexe Funktionen einer reellen Veränderlichen. Es folgen die Stetigkeit der Funktionen einer Veränderlichen und die damit zusammenhängenden Sätze. Das folgende Kapitel behandelt die Grenzwerte der Funktionen mehrerer Veränderlicher, ihre Stetigkeit, ihre Halbstetigkeit und ihre Transformationen. Weiterhin wird der Grenzwert und die Stetigkeit der Funktionen einer komplexen Veränderlichen behandelt. Anschließend werden die Begriffe der unendlich kleinen bzw. unendlich großen Werte eingeführt. In den folgenden Kapiteln kommen die Begriffe der Ableitung und des Differentials der Funktionen einer reellen Veränderlichen, sowie der Begriff der Stammfunktion zur Betrachtung. Im folgenden wird die Differentialrechnung auf Funktionen von mehreren Veränderlichen verallgemeinert (Extremenrechnung, Funktionenfolgen,

[5]) Á. Császár, Sur la structure des espaces de probabilité conditionnelle, *Acta Math. Acad. Sci. Hung.*, **6** (1955), 337—361.

Funktionenreihen, usw.). Es folgen die linearen Transformationen und die impliziten Funktionen, dann die Behandlung der funktionellen Abhängigkeit und Unabhängigkeit, die Betrachtung der Extreme mit Nebenbedingungen.

Die letzten fünf Kapitel des zweiten Bandes sind den Problemen der Orientierung in der Ebene und im Raum, der Untersuchung der Polygone, der Polyeder, der Kurven, Flächen und Körper gewidmet. Die Behandlung dieser Probleme ist hier wesentlich detaillierter, als in den meisten anderen Werken dieser Art.

Der dritte Band wird mit dem Integralbegriff Mengoli — Cauchy eingeleitet und dann geht er auf die Behandlung des Riemannschen Integrals über. Die Kurvenintegrale werden nach der Einführung der Rektifizierbarkeit von Kurven behandelt, dann befaßt sich der Band mit der Integration der linearen Differentialformen. Dem folgen einige Probleme der Berechnung von bestimmten Integralen. Anschließend wird die Laplace-Transformation untersucht. In drei Kapiteln werden die Theorie und Lösungsmethoden der gewöhnlichen Differentialgleichungen und Differentialgleichungssysteme, in einem Kapitel einige wichtige partielle Differentialgleichungen behandelt; es schließen sich die Grundtatsachen der trigonometrischen Reihen an. Der dritte Band behandelt noch das Maß der ebenen und räumlichen Mengen, Doppel-, dreifache, sowie Oberflächenintegrale.

Die Behandlungsmethode des Buches ist im allgemeinen detaillierter als in den meisten anderen Lehrbüchern. Didaktisch gesehen gelingt die Behandlung der Elemente der Funktionentheorie parallel mit der reellen Analysis ausgezeichnet. Der Gebrauch des Buches wird durch seine Übersichtlichkeit, durch die gute Gliederung der komplizierten Teile und durch die klare und exakte Ausdrucksweise erleichtert.

A. Kósa (Budapest)

Ákos Császár, Foundations of general topology, 380 pages, Pergamon Press, Oxford—London—New York—Paris, 1963.

Ákos Császár, Grundlagen der allgemeinen Topologie, 368 Seiten, Akadémiai Kiadó, Budapest, 1963.

Translations, by Mrs. K. Császár, of the first (French) edition (*Fondaments de la topologie générale*, reviewed in *Acta Sci. Math.*, **22** (1961), 321).

Some minor changes have been effected nearly in all chapters, three chapters (8, 15, 16) have been essentially rewritten, and four new chapters (17, 18, 19, 20) have been added. Chapters 17, 18 study the question of embedding syntopogenous spaces into cubes, i. e. into the Cartesian product of copies of the interval [0, 1], supplied with a suitable syntopogenous structure. For that reason the notion of the *rank function* of a syntopogenous structure has been introduced, the definition of which is based on the idea of weight of a topogenous order, which is a generalization of the notion of weight of a topological or proximity space. In chapter 19 the notion of a totally bounded syntopogenous space has been introduced, generalizing the idea of totally bounded metric and uniform spaces. Finally, in chapter 20, some supplementary remarks have been made concerning metrizable spaces.

The author expresses his indebtedness to the late J. Czipszer for his contributions to the completion of this book. Chapters 17. 18 and 19 are almost entirely due to Czipszer and contain his unpublished results.

L. Gehér (Szeged)

J. Horn—H. Wittich, Gewöhnliche Differentialgleichungen (Göschens Lehrbücherei, 1. Gruppe, Reine und angewandte Mathematik, Bd. 10), 6. Auflage, 275 Seiten, Berlin, Walter de Gruyter & Co., 1960.

As the number of the editions shows, this book is a well proved popular introduction into the theory of ordinary differential equations. The systematizing of the subject-matter, which is exemplary in didactical respect, the simplicity of the proofs, the great number of problems solved and examples given make this book suitable not only for students interested in differential equations and in their applications, but for teachers too.

In this edition there were made several changes. Especially the chapters considering differential equations in the complex domain and special functions satisfying some ordinary differential equations e. g. the Bessel differential equation were rewrited and extended. Also are improved the chapters dealing with the dependence of the solutions on a parameter and the initial conditions, and with the singularities of nonlinear differential equations.

The chapters are: elementary integration methods; existence of solutions; method of successive approximations; numerical and graphical approximation methods; linear differential equations; existence of solutions in the complex domain; linear differential equations of the second order; dependence of solutions on parameters and initial conditions; singularities of nonlinear differential equations; differential equations with periodic coefficients.

L. Pintér (Szeged)
Lajos

Martin Barner, Differential- und Integralrechnung, I, Grenzwertbegriff, Differentialrechnung (Sammlung Göschen, Band 86/86 a), 176 Seiten, Berlin, Walter de Gruyter & Co., 1961.

In diesem kleinen Buch werden die wichtigsten Fragen der Differentialrechnung der reellen Funktionen mit einer Variablen betrachtet: Eigenschaften der reellen Zahlen, Elemente der Mengenlehre, Begriff der reellen Funktion, Konvergenz der reellen Zahlenfolgen, Stetigkeit von Funktionen, Differenzierbarkeit der Funktionen (Mittelwertsätze, Taylorsche Formel und Reihe), axiomatische Definition der Logarithmusfunktion, der Exponentialfunktion und der Winkelfunktionen und ihre Eigenschaften.

Verf. geht von einem Axiomensystem für die reellen Zahlen aus, dann baut er die Theorie kompakt und lückenlos, in einer modernen Auffassung auf. Die einzelnen Axiome und Begriffe werden mit entsprechenden Aufgaben und Bemerkungen erläutert.

K. Tandori (Szeged)

E. F. Beckenbach and R. Bellman, Inequalities (Ergebnisse der Mathematik und ihrer Grenzgebiete, Neue Folge, Heft 30), XI+198 pages, Berlin—Göttingen—Heidelberg, Springer-Verlag, 1961; second revised printing, 1965.

Since the classic work on inequalities by HARDY, LITTLEWOOD and PÓLYA, appeared in 1934, the importance of inequalities in different branches of mathematics increased and changed in such a way that a new summary of the new discoveries, results and applications has become necessary. Let us only mention the theory of differential equations, the theory of games, and mathematical economics.

The book consists of five chapters. The first chapter contains the fundamental inequalities, beginning with the inequality between the arithmetic and geometric means, for which various proofs are given. Of course some inequalities and theorems only are mentioned, without proofs, but at the end of every chapter very detailed bibliographical notes are given. In chapter 2 a variety of interesting inequalities on matrices is discussed. After considering positive definite matrices and some simpler results, more intricate theorems are obtained by an integral representation of determinants, due to INGHAM and SIEGEL. In the second half of this chapter one finds results about positive matrices. These matrices play an important role e. g. in the study of computational algorithms for the numerical solution of partial differential equations. Chapter 3 deals with moment spaces and resonance theorems. Some of the paragraphs: Moments; Convexity; L^p-space; A result of F. RIESZ; Non-negative Trigonometric and Rational Polynomials; A Resonance Theorem of LANDAU; The Theory of Linear Inequalities; The Min-Max Theorem of VON NEUMANN. Chapter 4 deals with the following theme: "Given a set of functions $\{u\}$ satisfying certain side conditions, and an operator L that can be applied to the functions of this set, determine when the inequality $L(u)\geqq 0$ implies that $u\geqq 0$". The attention is focused upon ordinary and partial differential operators. In Chapter 5 inequalities for differential operators are considered. This chapter includes inequalities of B. SZ.-NAGY, HALPERIN, VON NEUMANN, CARLSON, WIRTINGER, etc. concerning e. g. inequalities between a function u and its derivatives $u^{(k)}$ and $u^{(n)}$.

Summing up: the book presents an enormous quantity of results from different branches of mathematics, lucidly arranged and with very useful hints to the literature.

The success of the book is marked by the fact that a second printing of it has been already necessary.

L. Pintér (Szeged)

INDEX — TARTALOM

— *Fodor, G.* and *Máté, A.* On the structure of set mappings and the existence of free sets 1
Freud, G. Approximation of continuous functions on compact metric space by linear methods 9
Fried, E. Representation of partially ordered groups 15
— *Leindler, L.* Über Konvergenz- und Summationseigenschaften von Haarschen Reihen 19
— *Gehér, L.* On completely continuous and uniformly bounded operators in l^p spaces 31
Fuchs, L. On partially ordered algebras. II 35
— *Гечег, Ф.* и *Пеак, И.* Автоматы с изоморфными полугруппами 43
— *Пеак, И.* Автоматы и полугруппы. II 49
— *Rédei, L.* Ein Überdeckungssatz für endliche abelsche Gruppen im Zusammenhang mit dem Hauptsatz von Hajós 55
Rényi, A. On certain representations of real numbers and on sequences of equivalent events 63
— *Tandori, K.* Über ein Problem von S. B. Stetschkin 75
— *Sz.-Nagy, B.* et *Foiaş, C.* Sur les contractions de l'espace de Hilbert. X. Contractions similaires à des transformations unitaires 79
Alexits, G. und *Králik, D.* Über die Approximation im starken Sinne 93
Hosszú, M. Über eine Verallgemeinerung der Distributivitätsgleichung 103
Jajte, R. General theory of summability. I. 107
— *Leindler, L.* Über verschiedene Konvergenzarten der trigonometrischen Reihen. II (Die Entwicklung der Ableitungen) 117
Brown, A., Halmos, P. R., Shields, A. L. Cesàro operators 125
Kovács, L. G. and *Neumann, B. H.* An embedding theorem for some countable groups 139
Kovács, L. G. and *Neumann, B. H.* On the existence of Baur-soluble groups of arbitrary height 143
Baumslag, G., Kovács, L. G., and *Neumann, B. H.* On products of normal subgroups 145
van Leeuwen, L. C. A. Über die zuläßigen Ideale in Szépschen Ringerweiterungen 149
Prohaska, L. Über die Existenz normaler Komplemente zu gewissen Hallgruppen 159
Seibt, H. Zur Quadratbildung von Polynomen 163
Weinert, H. J. Zur Theorie der Algebren und monomialen Ringe 171
Hildebrandt, S. Numerischer Wertebereich und normale Dilatationen 187
— *Sz.-Nagy, B.* Positive definite kernels generated by operator-valued analytic functions 191
— *Sz.-Nagy, B.* et *Foiaş, C.* Corrections et compléments aux Contractions IX. 193
Bibliographie 197

ACTA SCIENTIARUM MATHEMATICARUM

SZEGED (HUNGARIA), ARADI VÉRTANÚK TERE 1

Prix d'abonnement pour l'étranger $ 8.50. On peut s'abonner à l'entreprise de commerce des livres et journaux „Kultúra" (Budapest, I. Fő utca 32).

INDEX: 26024

65-5636 Szegedi Nyomda Vállalat

Felelős szerkesztő és kiadó: Szőkefalvi-Nagy Béla
A kézirat nyomdába érkezett: 1965. január hó
Megjelenés: 1965. június hó

Példányszám: 825. Terjedelem: 17,2 (A/5) ív
Készült monószedéssel, íves magasnyomással az MSZ 5601-54 és az MSZ 5602-55 szabvány szerint